

Db 81 AACGTGTCGACCCAGTCATCATCCATTTTCGAGGTGTCATGTCTCAGCTGCTGAGGTG 140
Qy 220 GAGGAGAGAACAGAGCTTTTAAATAACAAACATCTGGCTAAACATAGAGTGAATGATG 279
Db 141 GATGAAGTAAACAGAGATCATGGAGACCAACCTGTGGCTCAAGCAAAATCTGGAATGACTAC 200
Qy 280 AACTGAGGTGGAGACACTTCAGATTTTCGGCGGGGTCAAAGATTTAAGAGTGCACCCAC 339
Db 201 AAGCTGAAGTGGAAACCCCTCGACTATGGTGGGGGAGAGTTCAATGCGTGTCCCTGCAACG 260
Qy 340 AGACTATGGAACACAGACGTCCTTATGTACAAACAGCGCGGACGAAGGGTTTCGACGACG 399
Db 261 AAGATCTGGAAGCAGACATTTGCTGTATAACAAATGCTGTGGGATTTCAGGTGGAC 320
Qy 400 TATCCAAAGAGAGTGGTGGCGGACCAACGGCTCGTGTCTGTACGTGGCGCGCGGATC 459
Db 321 GACAAGACCAAGCCCTTACTCAAGTACACTGGGAGGTGACTTGGATACCTCCGGCCATC 380
Qy 460 TTCAGAGCACCTGCAAGATCGACATCACCTGGTTCCCTTCGACGACCAACGATGCGAG 519
Db 381 TTAAAGAGTCTCTGTAAATCGAGTGAACCTACTTCCCGTTGATTACCAAAACTGTACC 440
Qy 520 ATGAAGTTTGGCAGCTGCAATTATGATGTTATCAGTTGGATCTTACAACCTACAGGATGAA 579
Db 441 ATGAAGTTTCGTTCTCTGCTCCTACGATAAGCGCGAAATCGATCT---GGTCTGTATCGGC 497
Qy 580 GGGGGCGGAGATATAAGCAGTTTGTGCAGAAATGGCGAATGGGAGTTAATAGAGTCC 639
Db 498 TCTTCCATGAACCTCAAGGACTATTGGGAGAGCGCGAGTGGGCCATCATCAAGGCCCA 557
Qy 640 GGCAAGCGCAACAGAGATCTACTCAACTGTGTGCGGAGCGCATACATCGACATCAAGTTT 699
Db 558 GGCTACAAACAGACATCAAGTACACTGTGCGAGAGATCTACCCCGACATCAATAC 617
Qy 700 GCGGTGTGATCCGGAGGAAAACGCTCTACTACTTCTTCAATCTGATCGTGCCTCGGTG 759
Db 618 TCGCTGTACATCCGGCGCCTGCGCTTGTCTACACCATCAACCTCATCATCCCTGCGCTG 677
Qy 760 CTCATCGCTCCATGGTCTCTATTTGGGTTTCACTTGCCTCCAGACTCCGAGAAAGTTG 819
Db 678 CTCATCTCTCTCTCACTGTGCTGTCTTCTA CTGCTCCCTCCGACTCGCTGAGAGGTG 737
Qy 820 TCTTTAGTGTGACGATTAATCTACTGTCTGTGACGCTGTCTCAACATGGTGGCGGAGCG 879
Db 738 ACCCTGTGCAATTCTGTCTCTCTCCCTGACGGTGTCTCTCTGTTGATCACTGAGACC 797
Qy 880 ATGCCAGCGAGTGGAGCGCGCTTGTGCTCGGACCTACTTCAACTGCATCATGTTTC 939
Db 798 ATCCCTTCCACCTCGCTGCTCATCCCCCTGATTGGAGAGTACCTCTCTGTTTCAACCATGATT 857
Qy 940 ATGCTGGCTTCTCCGTCGTCCTCCACCATCTGATCTCTCACTTACCAACACCGCGCAGCA 999
Db 858 TTGTGAACCTTGCCATCGTCATCACCGTCTTCTGCTCAACGTGCACTACAGAACCCCG 917
Qy 1000 GACACTCAGAAATGATGATGATTCGTTGGGTGTTCTCTTTATTTGGCTGCGGTGGGTG 1059
Db 918 ACGACACACAAATGCCCTCATGGTGAAGACTGTATTTCTGAACCTGCTCCCGAGGTC 977
Qy 1060 CTGCGCATGTCAGGCC 1076
Db 978 ATGTTTCATGACGAGGCC 994

RESULT 12
AY402875
LOCUS Mus musculus CHRNA3 gene, VIRTUAL TRANSCRIPT, partial sequence, 1436 bp DNA linear GSS 15-DEC-2003
DEFINITION genomic survey sequence.
ACCESSION AY402875
VERSION AY402875.1 GI:39758858
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1436)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1436)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source Location/Qualifiers
1..1436
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1436
/gene="CHRNA3"
/locus_tag="HCM1369"
ORIGIN
Query Match 18.1%; Score 269.6; DB 29; Length 1436;
Best Local Similarity 55.1%; Pred. No. 1.1e-41;
Matches 549; Conservative 0; Mismatches 444; Indels 3; Gaps 1;
Qy 81 GGCACGCGTGGGTATCCACGAGAAGCGGTACTGACCACTATTGGACCACTACAACTG 140
Db 2 GGCACGCGCTCGAAGCTGAGCACCGCTGTTCCAGTACCTGTTTGAAGATTACAACGA 61
Qy 141 ACTGAGAGCGCCGTCGTCAACGAGAGCGACCCGCTGCGAGCTCTCTCTGGCCCTCAGCT 200
Db 62 GATCATCCGCGCGTGGCTAACTGTCGCCATCTCTGTCATCATCCAGTTTGGAGGTGTAT 121
Qy 201 CATGCGAGATCATCGAGTGGACGAGAACAGAGCTTTTAAACAACAATCTGGCTAAA 260
Db 122 GTCTCAGCTGGTGAAGTGGATGAAGTAAACAGATCATGGAACCAACCTGTGGCTGAA 181
Qy 261 ACTAGAGTGAATGATATGAATTTGAGGTGGAAACACTTTCAGATTTTCGGCGGGTCAAAGA 320
Db 182 GCAAACTGGAATGACTACAAGCTGAAATGGAAACCTCTGACTTACCAAGGGTGGAGTT 241
Qy 321 TTTAAGAGTGCACCCACAGACTATGGAACAGAGCTCTTATGTACAAACGCGCGGA 380
Db 242 CATGCGAGTCCCTGCGAGAGAAGATCTGGAAACAGACATCGTCTTTTACAACCAACGCGGA 301
Qy 381 CGAAGGCTTCGACAGCACGATATCCACGAACGTTGGTGGGAAACAACGCTCGTCTCT 440
Db 302 TGGGATTTCCAGTGGATGACAAACCAAGCTCTACTCAAGTACACAGAGAGTGC 361
Qy 441 GTACGTGCGCCCGGCATCTTCAAGAGACCTTCAAGAGACATCACTCTGGTTCCTCT 500
Db 362 TTGGATCCCTCCGGCCATCTTTAAGAGCTCATGCAAAATCATGATGTGACCTACTTCCGCTT 421
Qy 501 CGACGACCAACGATCGAGATGAAGTTTGGCACTGGACTTATGATGTTTATCAGTTGA 560
Db 422 TGACTACCAAACTGCACCATGAAGTTTCGGCTCTCTGCTCTACGACGAAGCAAGATCGA 481
Qy 561 TCTACAACTACAGATGAAGGGGGCGGAGATATAAGCAGTTTGTCCAGCAATGCGCAATG 620
Db 482 CCT---GGTCTCATTTGGCTCTCCATGAACCTCAAGGACTATTTGGGAAAGTGGCGAGT 538
Qy 621 GGAGTTAATAGGAGTCCCCGGCAAGCGCAACGAGATCTACTACAACTGTTTGTTCGGAGCC 680

Todd, M.A., Tanenbaum, D.M., Givello, D.R., Lu, F., Murphy, B.,
 Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.

FEATURES
 source 1..1436
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..>1436
 /gene="CHRNA3"
 /locus_tag="HCM1369"

ORIGIN
 Query Match 17.7%; Score 263.4; DB 29; Length 1436;
 Best Local Similarity 54.1%; Pred. No. 1.8e-40;
 Matches 529; Conservative 0; Mismatches 445; Indels 3; Gaps 1;
 QY 100 GAGAGGGGTACTGACACCACTATTGGACCACTACAACTGACTGGAGAGCGCGCTGCTC 159
 Db |||||
 QY 21 GAGCAGCGCTATTGAGCGGCTGTTTGAAGATTACAATGAGATCAATCGCGCTGTGGCC 80
 Db |||||
 QY 160 AAGCAGAGCAGCGCTGACGCTCTCCTCGGCTCAGCGCTCATGCGAGATCATCGAGCTG 219
 Db |||||
 QY 81 AAGTGCTGACCCAGTCATCATCTTNNNGTGTCCATGTCTNNNTGGTGAAGTG 140
 Db |||||
 QY 220 GACGAGAAGAACAGCTTTTAAACAAACATCTGGCTAAACCTAGAGTGAATGATG 279
 Db |||||
 QY 141 GATGAAGTAAACAGATCATGGAGACCACTGNNGTCAAGCAAACTCTGGAATGACTAC 200
 Db |||||
 QY 280 AACTGAGGTGGAACACTTCAGATTTCGGGGGTGCAAGATTTAAGATGCGACCCAC 339
 Db |||||
 QY 201 AAGCTGAAGTGAACCCCTCTGACTATGTTGGGGCAGAGTTCATGCGTGTCCCTGNNNN 260
 Db |||||
 QY 340 AGACTATGGAACACAGAGCTCTTATGTACAAACGCGCGAGCAAGGTTTCGACGACG 399
 Db |||||
 QY 261 NNNNCTGGAGCANNACATGTGCTGTATTAACAATGCTGTGGGATTTCCAGTGGAC 320
 Db |||||
 QY 400 TATCCAACAGCGTGTGTCGCGAAACACCGCTGCTGTCTGACGTGCGCGCCCGGCATC 459
 Db |||||
 QY 321 GACAAGACCAAGCTTACTCAAGTACACTGGGAGGTGACTTGGATACCTCCGGCCATC 380
 Db |||||
 QY 460 TTCAAGAGCCTCGAGATCGACATCACCTGTTCCCTTCGACGACCAACGATGCGAG 519
 Db |||||
 QY 381 TTTAAGAGCTCTCGCAAAATCGACGTGACCTACTTCCCGTTGATTACCAAAACTGTACC 440
 Db |||||
 QY 520 ATGAAGTTTGGCAGCTGACTTATGATGTTATCAGTTGATCTACAACCTACAGGATGAA 579
 Db |||||
 QY 441 ATGAAGTTGCTTCTGCTGCTAGTAAGCGGAATCAATCT--GGTCTGATCGGC 497
 Db |||||
 QY 580 GGGGGCGGAGATATAGCACTTTTGTCAAGATGGGAATGGAGTTTAATAGAGTCCCC 639
 Db |||||
 QY 498 TCTTCCATGAACCTCAAGGACTATTGGGAGAGCGGAGTGGGCCATCATCAAGGCCCA 557
 Db |||||
 QY 640 GGCAAGCGCAACGAGATCTACTCAACTGTGTCGGAGCGCATACATCGACATCACTTT 699
 Db |||||
 QY 558 GGCTACAAACAGACATCAAGTCAACTGCTNNNAGAGATCTACCCCGACATCAATAC 617
 Db |||||
 QY 700 GCGGTGCTGATCGGAGGAAAGCTCTACTCTTCTTCAATCTGATCGTCCCTGCGTG 759
 Db |||||
 QY 618 TCGCTGTACATCGGCGCTGCTTGTCTTACACCATCAACTCATCATCTCCCTGCGTG 677
 Db |||||
 QY 760 CTATCGCTCCATGCTTATGGGGTTTCACTTGCCTCCAGACTCCGAGAAAAGTTG 819
 Db |||||
 QY 678 CTCATCTCTCTCTCACTGTCTCTTCTTACCTGCTCCCTCCAGTCCGGTGAAGAGGT 737
 Db |||||
 QY 820 TCTTTAGGTGTGACGATTAATGCTGCTGACGCTGTCTCTCAACATGTTGGCGGAGACG 879
 Db |||||
 QY 738 ACCCTGTGAATTTCTGCTCTCTCTCTGACGGTGTCTCTCTGCTGATCATCTGAGACC 797
 Db |||||

QY 880 ATCCAGGACGCTCGAGCGCGCTGCTCGGACCTTACTTCAACTGCATCATGTTTC 939
 Db |||||
 QY 798 ATCCCTTCCACCTCGCTGCTGCTATCCCTGATGGAGAGTACCTCTGTTCCACATGATT 857
 Db |||||
 QY 940 ATGCTGCTTCTTCCCTGCTGCTCCACCATCTGATCTCAACTACCAACCGGACGCA 999
 Db |||||
 QY 858 TTTGTAACCTGCTGCTATCGCTGCTTCAACGCTGCTCAAGTGCCTACAGAACCCCG 917
 Db |||||
 QY 1000 GACACTCAGAAATGAGTATGATGCTGCTGCTGCTTCTTATTTGGCTGCGCTGGTG 1059
 Db |||||
 QY 918 ACGACACACACAATGCCCTCATGGGTGAAGACTGTATTTGAACCTGCTCCCGAGGTC 977
 Db |||||
 QY 1060 CTGCGCATGTCACGGCC 1076
 Db |||||
 QY 978 AUGTTCATGACAGGCC 994
 Db |||||

RESULT 15
 AK080475
 LOCUS
 DEFINITION
 Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
 enriched library, clone:A730038F14 product:NEURONAL NICOTINIC
 ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT, full insert sequence.
 AK080475
 AK080475.1 GI:26099260
 VERSION
 KEYWORDS
 HTc; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1
 CARNINCI, P. and HAYASHIZAKI, Y.
 TITLE
 High-efficiency full-length cDNA cloning
 JOURNAL
 Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE
 99279253
 PUBMED
 10349636
 2
 CARNINCI, P., SHIBATA, Y., HAYATSU, N., SUGAHARA, Y., SHIBATA, K.,
 ITOH, M., KONNO, H., OKAZAKI, Y., MURAMATSU, M. and HAYASHIZAKI, Y.
 TITLE
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL
 Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE
 20493374
 PUBMED
 11042159
 3
 SHIBATA, K., ITOH, M., AIZAWA, K., NEGAOKA, S., SASAKI, N., CARNINCI, P.,
 KONNO, H., AKIYAMA, J., NISHI, K., KITSUNAI, T., TASHIRO, H., ITOH, M.,
 SUMI, N., ICHI, Y., NAKAMURA, S., HAZAMA, M., NISHINE, T., HARADA, A.,
 YAMAMOTO, R., MATSUMOTO, H., SAKAGUCHI, S., IKEGAMI, T., KASHIWAGI, K.,
 FUJIWAKE, S., INOUE, K., TOGAWA, Y., IZAWA, M., OHARA, E., WATAHAKI, M.,
 YONEDA, Y., ISHIKAWA, T., OZAWA, K., TANAKA, T., MATSUURA, S., KAWAI, J.,
 OKAZAKI, Y., MURAMATSU, M., INOUE, Y., KIRA, A. and HAYASHIZAKI, Y.
 TITLE
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL
 Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE
 20530913
 PUBMED
 11076861
 4
 THE RIKEN GENOME EXPLORATION RESEARCH GROUP PHASE II TEAM AND THE
 FANTOM CONSORTIUM.
 TITLE
 Functional annotation of a full-length mouse cDNA collection
 JOURNAL
 Nature 409, 685-690 (2001)
 MEDLINE
 11076861
 5
 THE FANTOM CONSORTIUM AND THE RIKEN GENOME EXPLORATION RESEARCH
 GROUP PHASE I & II TEAM.
 TITLE
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 JOURNAL
 Nature 420, 563-573 (2002)
 MEDLINE
 12000000
 PUBMED
 12000000
 6
 (bases 1 to 2010)
 JOURNAL
 Nature 420, 563-573 (2002)
 MEDLINE
 12000000
 PUBMED
 12000000
 7
 ADACHI, J., AIZAWA, K., AKIMURA, T., ARAKAWA, T., BONO, H., CARNINCI, P.,
 FUKUDA, S., FURUNO, M., HANAGAKI, T., HARA, A., HASHIZUME, W.,
 HAYASHIZAKI, K., HAYATSU, N., HIRAMOTO, K., HIRAKAWA, T., HIROZANE, T.,

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:09:28 ; Search time 42.6667 Seconds

(without alignments)
3284.615 Million cell updates/sec

Title: US-09-303-232-4

Perfect score: 2665

Sequence: 1 MCGRRARRSHLAAGALLLL.....LFTIIATLAVLLSAPHIMWS 496

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2665	100.0	496	3 AAY50815	Aay50815 H. viresc
2	1850	69.4	770	3 AAY50814	Aay50814 D. melano
3	1803.5	67.7	501	3 AAY50816	Aay50816 H. viresc
4	1417.5	53.2	311	4 ABB63683	Abb63683 Drosophil
5	1267	47.5	502	2 AAW12368	Aaw12368 Neuronal
6	1253	47.0	498	4 ABB60432	Abb60432 Drosophil
7	1246	46.8	502	2 AAW44153	Aaw44153 Human neu
8	1246	46.8	502	2 AAW09025	Aaw09025 Neuronal
9	1246	46.8	502	3 AAB24088	Aab24088 Human PRO
10	1246	46.8	502	4 AAB82690	Aab82690 Nicotinic
11	1246	46.8	502	4 AAB50012	Aab50012 Wild-type
12	1246	46.8	502	5 ABB70492	Abb70492 Human neu
13	1246	46.8	502	5 ABB82435	Abb82435 Human neu
14	1246	46.8	502	7 ADA10874	Ada10874 Human neu
15	1246	46.8	502	7 ADD47051	Add47051 Human PRO
16	1246	46.8	502	7 ADE57310	Ades7310 Human PRO
17	1242	46.6	502	2 AAW69216	Aaw69216 Mutant hu
18	1240	46.5	502	4 AAB50015	Aab50015 Mutant hu
19	1236	46.4	502	4 AAB50016	Aab50016 Mutant hu
20	1232.5	46.2	502	7 ADE57308	Ades7308 Rat Prote
21	1232.5	46.2	502	7 ADD47049	Add47049 Rat Prote
22	1230	46.2	502	4 AAB50017	Aab50017 Mutant hu
23	1219.5	45.8	511	2 AAW12369	Aaw12369 Neuronal
24	1053.5	39.5	554	4 AAE12824	Aae12824 Caenorhab
25	1053.5	39.5	554	6 ABP96318	Abp96318 Caenorhab

ALIGNMENTS

RESULT 1

AAY50815

ID AAY50815 standard; protein; 496 AA.

XX AC AAY50815;

XX AC AAY50815;

DT 17-FEB-2000 (first entry)

XX H. virescens acetyl-choline receptor protein from clone Hva7-1.

DE DE Acetyl-choline receptor; nicotinic; insect; insecticide; screening;

KW neurotransmission; plant protection agent; conductance; ACHR.

XX Heliothis virescens.

XX DE19819829-A1.

PN 11-NOV-1999.

PD 04-MAY-1998; 98DE-01019829.

XX 04-MAY-1998; 98DE-01019829.

XX (FARB) BAYER AG.

XX Adamczewski M, Oellers N, Schulte T;

XX WPI; 2000-014207/02.

DR N-PSDB; AA224476.

XX New nucleic acid encoding a nicotinic acetylcholine receptor from

PT insects, used to identify potential insecticides.

XX Example 1a; Page 17-19; 26pp; German.

XX This invention describes a novel nucleic acid (NA) encoding a nicotinic

CC acetyl-choline receptor (I) from insects which can be used as an

CC insecticide. Inhibitors of (I) interfere with neurotransmission. (I)

CC (also vectors containing it, its regulatory regions, and antibodies

CC directed against (I)-encoded proteins) are used to screen for: (a) plant

CC protection agents that alter conductance of ACHR, potentially useful as

CC insecticides, or (b) genes which encode polypeptides that are involved in

CC formation of functionally related ACHR in insects. (I) are also used to

CC isolate and characterize the specified regulatory regions and for

CC recombinant production of (II). This sequence represents an acetyl-

XX choline receptor isolated from Heliothis virescens

XX Sequence 496 AA;

Aab50014 Chimeric
Aab50018 Mature ce
Abb60716 Drosophil
Aao17243 Modified
Ade57314 Human Pro
Ade57318 Human Pro
Aaw09022 Neuronal
Ada10857 Human neu
Aaw44155 Human neu
Aar73966 Alpha 2 s
Aaw09021 Neuronal
Abg61850 Prostace
Abg31800 Human neu
Abb82430 Human neu
Ada10855 Human neu
Adc71171 Human 205
Adc71169 Human 205
Add45584 Human Pro
Ade59169 Human Pro
Ada83810 Human CHR

QY 242 LYFFNLIVPCVLIIASALLGFTLPPDSGEKLSGLVTILLSLTVFLNVAETMPATSDAV 301
 Db 247 LYFFNLIVPCVLIIASALLGFTLPPDSGEKLSGLVTILLSLTVFLNVAETMPATSDAV 305
 QY 302 PLLG 305
 Db 307 PLLG 310

RESULT 5
 AAW12368
 ID AAW12368 standard; protein; 502 AA.
 XX AC AAW12368;
 XX 17-JUN-1997 (first entry)
 DT Neuronal alpha-bungarotoxin binding protein alpha 1; cholinergic;
 DE ligand binding; ion channel.
 KW Gallus sp.
 OS
 FH Key Location/Qualifiers
 FT Peptide 1..22
 FT Protein /label= Sig_peptide
 FT 23..502
 FT /label= Mat_protein
 XX US5599709-A.
 XX 04-FEB-1997.
 XX 28-SEP-1989; 89US-00413947.
 XX 28-SEP-1989; 89US-00413947.
 XX (SALK) SALK INST BIOLOGICAL STUDIES.
 PA Lindstrom JM, Schoepfer RD;
 PI WPI; 1997-118297/11.
 DR N-PSDB; AAT59196.
 XX New isolated neuronal alpha-bungarotoxin-binding protein DNA - used to
 PT screen cholinergic agents and other drugs which may affect ligand
 PT binding, ion channel or other activities of the protein.
 XX Example; Fig 2A-B; 18pp; English.
 XX The alpha subunit (AAW12368) and alpha2 subunit (AAW12369) amino acid
 CC sequences of chick neuronal alpha-bungarotoxin binding protein (ABBP)
 CC were deduced from newly isolated DNA molecules (AAT59196-97) obt'd. from
 CC an 18-day chick embryo cDNA library. ABBP subunits can be produced in
 CC recombinant host cells, pref. a bacterium, and used in the screening of
 CC cholinergic agents and other drugs that may affect the ligand binding.
 CC ion channel or other activity of intact ABBP subtypes. The ABBP alpha
 CC and alpha2 subunits can also be used to produce subunit peptides for use
 CC as immunogens for preparing antibodies to permit affinity purification of
 CC subtypes and their histological location
 XX Sequence 502 AA;
 SQ Query March 47.5%; Score 1267; DB 2; Length 502;
 Best Local Similarity 48.6%; Pred. NO. 2.5e-117;
 Matches 249; Conservative 83; Mismatches 144; Indels 36; Gaps 6;
 QY 1 MGGARRSHLAAPAGLLLLCLLWPRGACGVEKRLHLHLHDVNVLERPVWVESDPLQ 60
 Db 1 MGLRALMLWLLAAAGLV-----RESLQGEFQKLYKELLKNVPLERPVANDSQPLT 52
 QY 61 LSFGLTLMQIIDVDEKNQLITNIWLKLEWMDMLNWNWTSDFGGVKDLRVPPHRLWPKDV 120

Db 53 VYFTLSMQIMDVDEKNQVLTNIWLQMYWTDHYLQWNVSEYPCGVKNVRFDPDGLIWKPD 112
 QY 121 LMVNSADEGSDSYPTNVVVRNNGSCLYVPPGIFKSTCKIDITWFPDDQRCMKFGSWT 180
 Db 113 LLVNSADERFDATEFTNTLVNLSGHCQYLPPIGPKSSCYIDVRWFPFDVQKCNLFGSWT 172
 QY 181 YDGYQLDLQDEGGDISFVTNGEWELIGVPGKNEIYVNCPEPYVIDITFAVWIRK 240
 Db 173 YGWSLDLQGE---ADISGISNGEWDLVGIPKRTESFYECCKEYPDITFTVWRRR 229
 QY 241 TLYYFFNLIVPCVLIIASALLGFTLPPDSGEKLSGLVTILLSLTVFLNVAETMPATSDA 300
 Db 230 TLYYGLNLLIPCVLIIASALLGFTLPPDSGEKLSGLVTILLSLTVFLNVAEIMPATSDS 289
 QY 301 VPLLGTFFNCIMFVWASSVSTILINHYHRHATHEDMSDWIRCVFLVWVLRMSRPG 360
 Db 290 VPLIAQYFASFTMIIVGLSVVTVIVLQYHHHDPDGGKMPKWTVRVILLNWCWFLRMKRP 349
 QY 361 SATTPPPARVPPPPDLELRSSKSLLANVLID-----IDDFRHPQAOQ 404
 Db 350 E-----DKVRPACQHKQRCSSLSSMEMNTVSGQCSNGNMLYIGFRCLDGVHCTPTDS 403
 QY 405 PQCRYRGG--EENGAGLAAHSCFVDYELSLILKEIRVITDQMRKXDEADISRDKWF 462
 Db 404 GVICGRMTCSPTEENLLHSGHPSEG-DPDLAKILEEVRYIANRFRDQDEBEAICNEKWF 462
 QY 463 AAMVVDRLCLIIIFTLTIATLAVLSAPHIM 494
 Db 463 AASVVDRLCLMAFSVFTIICITIGILMSAPNFV 494

RESULT 6
 ABB60432
 ID ABB60432 standard; protein; 498 AA.
 XX AC ABB60432;
 XX 26-MAR-2002 (first entry)
 DT Drosophila melanogaster polypeptide SEQ ID NO 8088.
 DE Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 KW Drosophila melanogaster.
 OS Drosophila melanogaster.
 XX WO200171042-A2.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US009231.
 XX 23-MAR-2000; 2000US-0191637P.
 XX 11-JUL-2000; 2000US-00614150.
 PA (PEKE) PE CORP NY.
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 XX N-PSDB; ABL04535.
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX Disclosure; SEQ ID NO 8088; 21pp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of


```
CC Alzheimer's disease, addiction to nicotine or schizophrenia
XX
SQ Sequence 502 AA;

Query Match 46.8%; Score 1246; DB 4; Length 502;
Best Local Similarity 46.0%; Pred. No. 3.1e-115;
Matches 240; Conservative 86; Mismatches 126; Indels 70; Gaps 6;

QY 12 APAGLLLLCLLWPRGARGVHEKRLHLLHLDHYNVLERPVVNESDPLQLSGLTMOII 71
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
4 SPGGVWIALAASLLHVSQGEFQKLYKELVKNYNPLRPVANDSQPTVTVFSLSLQIM 63
QY 72 DVEKKNOLLITNIWLEWNDMLRWNTSDFGVKDLRVPVPHRLWKPDVLMYNSADEGFD 131
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
64 DVEKKNQVLTNIWQMSWTDHYLQWNVSEYGVKTVRFDPGQIWKPDILLYNSADERFD 123
QY 132 STYPTNVVRNNGSLYVPPGIFKSTCKIDITWPFDDQRCMKFGSWTYGYQLDLQ 191
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
124 ATFTNTVLNVSNGHCQYLPPIFKSSCYIDVRWFPFDVQHCKLKFSGWSGLDLMQ 183
QY 192 DEGGDITSSFTNGEWELIGVPGKRNETIYNCCPEPIDITFAVVIREKTLIYFNLIVP 251
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
184 E---ADISGYTPNGEWDLVGIPGRSERFYECCKEYEDVTFVTMRRRTIYYGLNLLIP 240
QY 252 CVLIASALLGFTLPDPSGEKLSGLVTILLSLTVFLNVAETMPATSDAVLLGTYNFI 311
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
241 CVLISALALLVFLPADSGEKISLGITVLLSLTVFMLLVAEIMPATSDSVPLIAQYFAST 300
QY 312 MFMVASSVSTILINHYHRHADTHEMSDWTRCVFLYWLFWLWLRMSRPG----- 360
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
301 MIVGLSVVTVIVLQYHHDPDGCKMPKWTIRVILLANCAWFLRMKRPGEKVRPACQHK 360
QY 361 -----SATTTPPPAR-----VPPPDLELRSSKSLILANVLD 392
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
361 QRCSLASVMSAVAPPPASNGNLLYIGFRGLDGVHCVPTPDSGVVCGRMACSPH----- 416
QY 393 IDDDFRHPQAQQPCCRYYRGEENGAGLAHSCFGVDYELSLILKEIRVITDQMRKDE 452
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
417 -DEHLH-----GQPPPEG-----DPLAKILEEVRYIANRFRQDE 452
QY 453 DADISRDKWKAAMVDRCLLIIFTIITLAVLLSAPHIM 494
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
453 SEAVCSEWKFAACVVDRLCLMAFSVFTICTIGILMSAPNFV 494

RESULT 12
ABG70492
ID ABG70492 standard; protein; 502 AA.
XX
AC ABG70492;
XX
XX 06-DEC-2002 (first entry)
XX
DE Human neuronal nicotinic acetylcholine receptor alpha 7 subunit.
XX
KW Human; neuronal nicotinic acetylcholine receptor; nNACHr; receptor;
XX
KW ion flux; alpha 7 subunit.
XX
OS Homo sapiens.
XX
```

```
CC Alzheimer's disease, addiction to nicotine or schizophrenia
XX
SQ Sequence 502 AA;

Query Match 46.8%; Score 1246; DB 4; Length 502;
Best Local Similarity 46.0%; Pred. No. 3.1e-115;
Matches 240; Conservative 86; Mismatches 126; Indels 70; Gaps 6;

QY 12 APAGLLLLCLLWPRGARGVHEKRLHLLHLDHYNVLERPVVNESDPLQLSGLTMOII 71
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
4 SPGGVWIALAASLLHVSQGEFQKLYKELVKNYNPLRPVANDSQPTVTVFSLSLQIM 63
QY 72 DVEKKNOLLITNIWLEWNDMLRWNTSDFGVKDLRVPVPHRLWKPDVLMYNSADEGFD 131
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
64 DVEKKNQVLTNIWQMSWTDHYLQWNVSEYGVKTVRFDPGQIWKPDILLYNSADERFD 123
QY 132 STYPTNVVRNNGSLYVPPGIFKSTCKIDITWPFDDQRCMKFGSWTYGYQLDLQ 191
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
124 ATFTNTVLNVSNGHCQYLPPIFKSSCYIDVRWFPFDVQHCKLKFSGWSGLDLMQ 183
QY 192 DEGGDITSSFTNGEWELIGVPGKRNETIYNCCPEPIDITFAVVIREKTLIYFNLIVP 251
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
184 E---ADISGYTPNGEWDLVGIPGRSERFYECCKEYEDVTFVTMRRRTIYYGLNLLIP 240
QY 252 CVLIASALLGFTLPDPSGEKLSGLVTILLSLTVFLNVAETMPATSDAVLLGTYNFI 311
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
241 CVLISALALLVFLPADSGEKISLGITVLLSLTVFMLLVAEIMPATSDSVPLIAQYFAST 300
QY 312 MFMVASSVSTILINHYHRHADTHEMSDWTRCVFLYWLFWLWLRMSRPG----- 360
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
301 MIVGLSVVTVIVLQYHHDPDGCKMPKWTIRVILLANCAWFLRMKRPGEKVRPACQHK 360
QY 361 -----SATTTPPPAR-----VPPPDLELRSSKSLILANVLD 392
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
361 QRCSLASVMSAVAPPPASNGNLLYIGFRGLDGVHCVPTPDSGVVCGRMACSPH----- 416
QY 393 IDDDFRHPQAQQPCCRYYRGEENGAGLAHSCFGVDYELSLILKEIRVITDQMRKDE 452
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
417 -DEHLH-----GQPPPEG-----DPLAKILEEVRYIANRFRQDE 452
QY 453 DADISRDKWKAAMVDRCLLIIFTIITLAVLLSAPHIM 494
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
453 SEAVCSEWKFAACVVDRLCLMAFSVFTICTIGILMSAPNFV 494

RESULT 11
AAB50012
ID AAB50012 standard; protein; 502 AA.
XX
AC AAB50012;
XX
XX 14-MAR-2001 (first entry)
XX
DE Wild-type human alpha7 ligand gated ion channel.
XX
KW Human; alpha7 nicotinic acetylcholine gated ion channel;
XX
KW 5-hydroxytryptamine; 5-HT3; calcium ion conductance.
XX
OS Homo sapiens.
XX
PN WO200073431-A2.
XX
PD 07-DEC-2000.
XX
PF 25-MAY-2000; 2000WO-US011862.
XX
PR 27-MAY-1999; 99US-0136174P.
XX
PA (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI Groppi VE, Wolfe ML, Berkenpas MB;
XX
XX WPI; 2001-061524/07.
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2004, 06:45:49 ; Search time 5459.09 Seconds
(without alignments)
3938.047 Million cell updates/sec

Title: US-09-303-232-4

Perfect score: 2665

Sequence: 1 MGRARRSHAPAGLLLLL.....LFTIATLAVLLSAPHIMVS 496

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09303232/runat_07052004_101110_23883/app_query.fasta_1.2261
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAFSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09303232 @CGN 1.1 8225 @runat_07052004_101110_23883 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.em.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rod.*
36: em.htg_man.*
37: em.htg_vrt.*
38: em.sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2665	100.0	3629	3	AF143846 Heliothis
2	2665	100.0	3700	6	AX009612 Sequence
3	2665	100.0	3701	6	E58347 Nucleic aci
4	2064.5	77.5	2907	3	AF272778 Drosophil
5	2047	76.8	1683	3	DMES54210
6	1912	71.7	2834	3	AY036613 Drosophil
7	1850	69.4	2886	6	E58346 Nucleic aci
8	1850	69.4	2886	6	AX009610 Sequence
9	1809	67.9	3029	3	AF143847 Heliothis
10	1809	67.9	3109	6	E58348 Nucleic aci
11	1809	67.9	3109	6	AX009614 Sequence
12	1795.5	67.4	2023	3	AF321446 Drosophil
13	1792.5	67.3	2023	3	AF321447 Drosophil
14	1791.5	67.2	1665	3	DMES54209
15	1791.5	67.2	1699	3	BT011147 Drosophil
16	1790.5	67.2	2023	3	AF321445 Drosophil
17	1774	66.6	2068	3	AF321448 Drosophil
18	1769	66.4	2110	3	AF321449 Drosophil
19	1386.5	52.0	2116	3	AY036614 Drosophil
20	1274.5	47.8	2037	5	GNNARA7A
21	1272.5	47.7	1719	5	AY247962 Dario rer
22	1267	47.5	2696	5	GGANAREC
23	1253	47.0	1668	5	AY298752 Takifugu
24	1250	46.9	1848	10	MUSNARS
25	1248	46.8	2088	10	AF225980
26	1247	46.8	1876	6	AR055255 Sequence
27	1247	46.8	1876	6	AR071403 Sequence
28	1247	46.8	1876	6	AR173187 Sequence
29	1247	46.8	1876	6	AR224030 Sequence
30	1247	46.8	1876	6	AR282833 Sequence
31	1247	46.8	1876	6	AX179088 Sequence
32	1247	46.8	1876	9	HSU62436 Human nicot
33	1246	46.8	1509	6	AX054567 Sequence
34	1246	46.8	1509	9	HSNACHRA7
35	1242	46.6	1590	6	BD021850 Sequence
36	1242	46.6	1590	6	BD023656 Variant h
37	1242	46.6	1916	9	AF385585 Homo sapi
38	1242	46.6	1977	9	HSU40583 Human alpha
39	1242	46.6	2106	10	RATNARAD
40	1241	46.6	1559	9	HUMAYNAR
41	1241	46.6	2087	9	HSARA7A
42	1240	46.5	1509	6	AX054575 Sequence
43	1239.5	46.5	1461	5	AY299465 Takifugu
44	1237.5	46.4	2107	9	AF486623 Macaca mu
45	1237.5	46.4	3030	10	SS3987 nicotinic r

ALIGNMENTS

RESULT 1

AF143846
 LOCUS 3629 bp mRNA linear INV 27-MAY-1999
 DEFINITION Heliothis virescens putative nicotinic acetylcholine receptor alpha
 7-1 subunit mRNA, complete cds.
 ACCESSION AF143846
 VERSION AF143846.1 GI:4895004
 KEYWORDS
 SOURCE
 ORGANISM
 Heliothis virescens (tobacco budworm)
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 Noctuoidea; Noctuidae; Heliothinae; Heliothis.
 1 (bases 1 to 3629)
 SCULTE, T., Oellers, N. and Adamczewski, M.
 Putative alpha subunits of insect nicotinic acetylcholine receptors
 more similar to vertebrate alpha 7 subunits and C. elegans Ce21
 than to other insect nicotinic acetylcholine receptor alpha
 subunits
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 3629)
 AUTHORS Schulte, T., Oellers, N. and Adamczewski, M.
 TITLE Direct Submission
 JOURNAL Submitted (19-APR-1999) ZF-BTB, Bayer AG, Bldg. Q 18,, Leverkusen
 51368, Germany

FEATURES

Location/Qualifiers
 1..3629
 /organism="Heliothis virescens"
 /mol_type="mRNA"
 /db_xref="taxon:7102"
 335..1825
 /note="hvnachra-7-1"
 /codon_start=1
 /product="putative nicotinic acetylcholine receptor alpha
 7-1 subunit"
 /protein_id="AAD32697.1"
 /db_xref="GI:4895005"
 /translation="MGRARRSLAAPAGLLLLCLLPWPGARCYHEKRLHLLDH
 YNLEFVNESDPQLSFGILTMQIIDVEKNQLITNIWLKLEWDMNLWNSTDF
 GVKVLDVPHRLWKPVLVMSADGFDSTPTNVVVRNNGSLVYPGIFKSTCKI
 DIWFPDQRCBKMFGSWTVYDGLDQLQDEGGDISSFTVINGEWELIGVPGKNE
 IYNCCPEYDITFAVIVRKTLVYFNLVPCVLASMALLOFTLPPSGEKLSLG
 VTLLSLTVFLNVAETMPATSDAPVLGYFNCIMFWASSVVSTILINLHRRHAD
 THEMDSIRCVFLVFLWLRMSRPGSATPPPARVPPPDLELRERSSKSLANVLD
 IDDFRHPQAPQCCYRGEENGAGLAHSCFQVDFYELSLIKELRIVITDQMRKD
 DEDADISRWKFAAMVVDRLIIFLTFTIATLAVLLSAPHIMVS"

ORIGIN

Alignment Scores:
 Pred. No.: 5,82e-247 Length: 3629
 Score: 2665.00 Matches: 496
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0
 US-09-303-232-4 (1-496) x AF143846 (1-3629)
 QY 1 MetGlyGlyArgAlaArgSerHisLeuAlaAlaProAlaGlyLeuLeuLeuLeu 20
 Db 335 ATGGCGCGCGCGCGCGCGCTGCACCTTGGCGCGCGCGCGCGCTGCTGCTGCTG 394
 QY 21 CysLeuLeuTrpProArgGlyAlaArgCysGlyTyrHisGlyLysArgLeuHisHis 40
 Db 395 TGCGTGTCTGGCGCGCGCGCGCGCTGGCGGTACCACGAGAGCGGCTACTGCACAC 454
 QY 41 LeuLeuAspHisTyrAsnValLeuAluArgProValValAsnGluSerAspProLeuGln 60
 Db 455 CTATTGGACCACTACACAGTACTGGAGAGCGCGCTGCTCAACGAGACGCCCGCTGCG 514
 QY 61 LeuSerPheGlyLeuThrLeuMetGlnIleLeuAspValAspGluLysAsnGlnLeuLeu 80
 Db 515 CTCTCTCTGGCGCTCACGCTCATGCATCATGCAGTGGACGAGAGACCAACAGCTTTTA 574

QY 81 IleThrAsnIleTyrLeuLysLeuGluTrpAsnAspMetAsnLeuArgTyrAsnThrSer 100
 Db 575 ATACAAACATCTGGCTAAACTAGAGTGGATATGATGAATGAATCTGAGGTGGACACTTCA 634
 QY 101 AspPheGlyGlyValLysAspLeuArgValProProHisArgLeuTyrLysProAspVal 120
 Db 635 GATTTCGGCGGGGTCAAAGATTAAAGAGTGCACCCACAGACTATGCAAAACGACGCTC 694
 QY 121 LeuMetTyrAsnSerAlaAspGluGlyPheAspSerThrTyrProThrAsnValValVal 140
 Db 695 CTTATGTACACACAGCGCGGACGAAGGTTTCACACAGCAGTATCCAAACGAAACGTGGTGT 754
 QY 141 ArgAsnAsnGlySerCysLeuTyrValProProGlyLysPheLysSerThrCysLysIle 160
 Db 755 CGGAACAAACGCTGCTGCTGTACGTGCCGCCCGGCATCTTCAAGAGCAGCTGCAAGATC 814
 QY 161 AspIleThrTrpPheProPheAspAspGlnArgCysGluMetLysPheGlySerTrpThr 180
 Db 815 GACATCACTGGTTCCCTTCGACGACCAACAGATCGAGATGAAGTTTGGCAGCTGGACT 874
 QY 181 TyrAspGlyTyrGlnLeuAspLeuGlnLeuAspGluGlyGlyAspIleSerSer 200
 Db 875 TATGATGTTATCAGTTGGATCTCAACTACAGGATGAAGGGCGGAGATATAGCAGT 934
 QY 201 PheValThrAsnGlyGluTrpGluLeuIleGlyValProGlyLysArgAsnGluIleTyr 220
 Db 935 TTTGTCAGGAATGCGGAATGGGAGTTAATAGGAGTCCCGGCAAGCGCAACGAGATCTAC 994
 QY 221 TyrAsnCysCysProGluProTyrIleAspIleThrPheAlaValIleArgArgLys 240
 Db 995 TACAACCTGTTGTCGGGAGCCATACATCGACATCACTGTTGGTGGTGGTATCCGAGGAAA 1054
 QY 241 ThrLeuTyrTyrPhePheAsnLeuIleValProCysValLeuIleAlaSerMetAlaLeu 260
 Db 1055 ACGCTCTACTACTTCTTCAATCTGATCGTGCCTCGTGTCTCATCGCTCCATGGCTCTA 1114
 QY 261 LeuGlyPheThrLeuProAspSerGlyGluLysLeuSerLeuGlyValThrIleLeu 280
 Db 1115 TTGGGGTTACCTTGCCTCCAGACTCCGAGAGAAAGTTGCTTTAGGTGTGACCATATTA 1174
 QY 281 LeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAla 300
 Db 1175 CTGTCGTGTAGCGGTTCCTCAACATGTTGGCGGAGACGATGCCAGCGACGCTCGACGCC 1234
 QY 301 ValProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAlaSerSerValVal 320
 Db 1235 GTGCCCTTGTTCGCGACCTCTTCACTGCATCATGTTCACTGTTGGTGTCTCTCGTCTG 1294
 QY 321 SerThrIleLeuIleLeuAsnTyrHisHisAlaAspThrHisGluMetSerAsp 340
 Db 1295 TCCACCATACCTGATCTCACTACCAACCGCGCACGACAGACACTCAGAAATGAGTGAT 1354
 QY 341 TrpIleArgCysValPheLeuTyrTrpLeuProTrpValLeuArgMetSerArgProGly 360
 Db 1355 TGGATTCTGTTGCGGTTCCTTTATTTGCTCGCGGTGCTGCGCATGTTCACGCGCGCGC 1414
 QY 361 SerAlaThrThrProProAlaArgValProProProProAspLeuGluLeuArgGlu 380
 Db 1415 TCGCGNACGACGCGCGCGCGCGGTACTCTCCGCCCGCGGACCTGGAGATCGCGCGAG 1474
 QY 381 ArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHisPro 400
 Db 1475 CGCTCTCTCAAGTCGCTCTTAGCGAACGCTGCTGACATCGATGACGACTTCCGCGCACCG 1534
 QY 401 GluAlaGlnGlnProGlnCysCysArgTyrTyrArgGlyGlyGluGluAsnGlyValGly 420
 Db 1535 CAAGCGCAGCAGCGCAATCTCCGATACTCAGGGGGGGTGGAGGAGATGGCGCGCGG 1594
 QY 421 LeuAlaAlaHisSerCysPheGlyValAspTyrGluLeuSerLeuIleLeuLysGluIle 440
 Db 1595 TTGGCGCGCAGTAGTGTCTCGGTGTGACATGAGCTCTCCCTCATCTTCGAGAGAGATT 1654
 QY 441 ArgValIleThrAspGlnMetArgLysAspGluAspAlaAspIleSerArgAspTrp 460

```
Db      1655 AGAGTCATCACAGATCAGATGCGCAGGACGACGAGAGATTCGGCGGACTGG 1714
Qy      461 LysPheAlaAlaMetValValAspArgLeuCysLeuIlePheThrLeuPheThrIle 480
Db      1715 AAGTTGCGCGCCATGTCGTGGACAGACTGTCCTTATTATCTTTACCCCTGTTCAACAATC 1774
Qy      481 IleAlaThrLeuAlaValLeuLeuSerAlaProHisIleMetValSer 496
Db      1775 ATGCGCACGCTAGCCGTGCTGTGTCGCGCCACACATCATGCTGTGTCG 1822

RESULT 2
AX009612 LOCUS AX009612 3700 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 3 from Patent EP0962528.
ACCESSION AX009612
VERSION AX009612.1 GI:9996844
KEYWORDS Heliothis virescens (tobacco budworm)
SOURCE Heliothis virescens
ORGANISM Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE 1
AUTHORS Adamczewski, M.D., Schulte, T.D. and Oellers, N.D.
TITLE Nucleic acids encoding acetylcholin-receptor subunits from insects
JOURNAL Patent: EP 0962528-A 3 08-DEC-1999;
BAYER AG (DE)
FEATURES
source 1..3700
/organism="Heliothis virescens"
/mol_type="unassigned DNA"
/db_xref="taxon:7102"
335..1825
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC07500.1"
/db_xref="GI:9996845"
/db_xref="RENTREMBL:CAC07500"
/translators="MGGRRARRSHLAAPAGLLLLCLLWPRGARGYHEKRLHLLHLDH
YVLERPVNESPLOLSFGLTLMQIIDVDKKNOLLITNWLKLEWDMNLRWNTSDP
GGVKDLRVPRLKPDVLVMSADGDFSTYPTNVVNRNGSLVYVPGIFKSTCKI
DIWFPPDORCEKGSWYDGYQLDLQDGGGDISSFVFTNGEWELGVPKQNE
IYNCCPEPIDITFAVVRKTLTYFFNLDIVPCVLIASWALLGFTLPPDSGKLSIG
VTLLSLTFLVNWVAETMPATSDAVPLLTGFNFCIMPVASSVSTLILILYHHRHAD
THEMSWIRCVFTYLPWLRMRSGASTPPPARVPPDLELRERSKSLANLVD
IDDDFRHQAOQCCRYRGGENGAGLAHSCFVVDYELSLILKEIRVITDQMRKD
DEDADISRDKWFAAMVVDRLCLIFLTITLAVLSAPHIMVS"

Alignment Scores:
Pred. No.: 5,99e-247 Length: 3700
Score: 2665.00 Matches: 496
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-303-232-4 (1-496) x AX009612 (1-3700)

Qy      1 MetGlyArgAlaArgArgSerHisLeuAlaAlaProAlaGlyLeuLeuLeuLeu 20
Db      335 ATGGGGCGGGCGGGCGCGCTCGACATTGGCGCGCGCGCGCGCGCTGCTGCTGCTG 394
Qy      21 CysLeuLeuTrpProArgGlyAlaArgCysGlyTyrrHisGluLysArgLeuLeuHisHis 40
Db      395 TGCCTGCTCTGGCGGCGGGGCGGCGCTGCGGGGTACACAGAGAGCGGCTACTGCACCAC 454
Qy      41 LeuLeuAspHisTyrAsnValLeuGluArgProValValAsnGluSerAspProLeuGln 60
Db      455 CTATTGGACCACTACAACGTACTGAGAGGCGCGCTCGTCAACGAGAGCGACCCGCTGCAG 514
Qy      61 LeuSerPheGlyLeuThrLeuMetGlnIleIleAlaValAspGluLysAsnGlnLeuLeu 80
```

```
Db      515 CTCCTCTTCGGCTCAGCTCATGCGATCATCGAGTGGACGAGAACACGACTTTTA 574
Qy      81 IleThrAsnIleTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSer 100
Db      575 ATAACAAACATCTGGCTAAAACTAGAGTGGAAATGATATGAACCTTGAGGTGGAACACTTCA 634
Qy      101 AspPheGlyGlyValLysAspLeuArgValProProHisArgLeuTrpLysProAspVal 120
Db      635 GATTTTCGGCGGGGTCAAGATTTTAAAGATGCGCCACCCACAGATATGGAACACGAGCTC 694
Qy      121 LeuMetTyrAsnSerAlaAspGluGlyPheAspSerThrTyrProThrAsnValValVal 140
Db      695 CTTATGTCACACAGCGCGGACGAGGGTTTCGACAGACCGTATCCACGACAGCTGGTGGTG 754
Qy      141 ArgAsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSerThrCysLysIle 160
Db      755 CGGAACAACGGCTCGTGTCTGTACGTGCGCGCGCGCATCTTCAAGAGCACCTCGACGATC 814
Qy      161 AspIleThrTrpPheProPheAspAspGlnArgCysGluMetLysPheGlySerTrpThr 180
Db      815 GACATCACCTGGTTCCTTCGACGACCAACGATGCGAGATGAAGTITTCGACGCTGACT 874
Qy      181 TyrAspGlyTyrGlnLeuAspLeuGlnLeuAspGluGlyGlyGlyAspIleSerSer 200
Db      875 TATGATGGTTATCAGTTGGATCTACAACTACAGATGAGGGGCGGAGATATAAGCAGT 934
Qy      201 PheValThrAsnGlyGluTrpGluLeuIleGlyValProGlyLysArgAsnGluIleTyr 220
Db      935 TTTGTACGAATGGCAATGGGAGTTAATAGGAGTCCCGCGCAAGCGCAACGAGATCTAC 994
Qy      221 TyrAsnCysCysProGluProTyrIleAspIleThrPheAlaValValIleArgAspLys 240
Db      995 TACAACCTGTTGTCGGAGGCATACATCGACATCACGTTTCGGGTGGTGAATCCGAGGAAA 1054
Qy      241 ThrLeuTyrTyrPhePheAsnLeuIleValProCysValLeuIleAlaSerMetAlaLeu 260
Db      1055 ACCTCTACTACTCTTCATCTGATCGTCCCTGGCTCATCGCTCCATCGCTCTA 1114
Qy      261 LeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeuGlyValThrIleLeu 280
Db      1115 TTGGGGTTACCTTGGCTCCAGACTCCGAGAAAAGTTGTCTTATAGGTGTGACGATATTA 1174
Qy      281 LeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAla 300
Db      1175 CTGTCGTGTAGCGGTGTTCTCAACATGTCGGGAGACGATCCGACGAGCTCGACGCC 1234
Qy      301 ValProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAlaSerSerValVal 320
Db      1235 GTGCCCTTGTTCGGCACCTTCACTCACTGCATCATGTTTCATGGTGGCTTCTCTCGTCGTC 1294
Qy      321 SerThrIleLeuLeuLeuAsnTyrHisHisArgHisAlaAspThrHisGluMetSerAsp 340
Db      1295 TCCACCATCTAGTTCCTCAACTACCAACCGGACGACGACACTCAGAAATGAGTGAT 1354
Qy      341 TrpIleArgCysValPheLeuTyrTrpLeuProTrpValLeuArgMetSerArgProGly 360
Db      1355 TGGATTCTGTCGTGTCTCTTTATTTGGCTCCGCGGGTGTCTGCCCATGTACGCCGCCGCG 1414
Qy      361 SerAlaThrThrProProAlaArgValProProProProAspLeuLeuLeuLeuLeu 380
Db      1415 TCGGCGAGACGCGCGCGCGCGGTACCTCCGCGCGCGGAGCTGGAGTGGCGGAG 1474
Qy      381 ArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspAspPheArgHisPro 400
Db      1475 CGCTCTCTCCAAGTCGTCTCTAGCGAAGCGTCTGCATCGATCGACATCGACGCTTCCGCGCACCCG 1534
Qy      401 GlnAlaGlnProGlnCysArgTyrTyrArgGlyGlyGluGluAsnGlyValagly 420
Db      1535 CAACGCGAGACGCGCGCAATCTGCCGATCTACAGGGGGGGTGGAGAGATGGCGCGGGG 1594
Qy      421 LeuAlaAlaHisSerCysPheGlyValAspTyrGluLeuSerLeuLeuLeuLeuLeu 440
```



```

Db 1595 TTGGCGCGCAGATTGCTCGGTGTCGACTACGAGCTCTCCCTCATTTCTGAAGAGATT 1654
Qy 441 ArgValIleThrAspGluMetArgLysAspAspGluAspAlaAspIleSerArgAspTrp 460
Db 1655 AGAGTCATCAGATCAGATCGCAGGACGACGACGACGACGACATTCGCGCGACTGG 1714
Qy 461 LysPheAlaMetValValAspArgLysLeuLeuIlePheThrLeuPheThrIle 480
Db 1715 AAGTTCCGCGCATGGTGTGGACAGACTGTGGCTTATATCTATTACCTGTTCAATC 1774
Qy 481 IleAlaThrLeuAlaValLeuLeuSerAlaProHisIleMetValSer 496
Db 1775 ATCGCCACGCTAGCGTGTGTGTGTCGCGGCACACATCATGTGTGTCG 1822

RESULT 3
E58347
LOCUS E58347 3701 bp DNA linear PAT 18-JUN-2001
DEFINITION Nucleic acid encoding insect acetyl choline receptor subunit.
ACCESSION E58347
VERSION E58347.1 GI:13019346
KEYWORDS JP 2000023680-A/2.
SOURCE Heliothis virescens (tobacco budworm)
ORGANISM Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
1 (bases 1 to 3701)
Martin, A., Nadja, E. and Thomas, S.
Nucleic acid encoding insect acetyl choline receptor subunit
Patent: JP 2000023680-A 2 25-JAN-2000;
BAYER AG
COMMENT
OS Heliothis virescens
PN JP 2000023680-A/2
PD 25-JAN-2000
PE 26-APR-1999 JP 1999118159
PR 04-MAY-1998 DE 19819829.9
PI MARTIN ADAMTUSUEUSUKI, NADJA ERASU, THOMAS SCHULTE PC
C12N15/09, A01K67/033, C07K14/705, C07K16/28, C12N1/21, C12N5/10, PC
C12Q1/68,
PC G01N33/15, G01N33/50// (C12N1/21, C12R1/19), C12N15/00, C12N5/00 CC

FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..3701
/organism="Heliothis virescens"
/mol_type="genomic DNA"
/db_xref="taxon:7102"

ORIGIN
Alignment Scores:
Pred. No.: 5.99e-247 Length: 3701
Score: 2685.00 Matches: 496
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-303-232-4 (1-496) x E58347 (1-3701)

Qy 1 MetGlyClyArgAlaArgArgSerHisLeuAlaAlaProAlaGlyLeuLeuLeuLeu 20
Db 335 ATGGGGCGGGCGGGCGCGCGCTGCACTTGGCGCGCGCGCGCGCTGCTGCTGCTG 394
Qy 21 CysLeuLeuTrpProArgGlyAlaArgCysGlyTyrrHisGluLysArgLeuLeuHis 40
Db 395 TGCTGCTCTGGCCAGGGGGGACGCTGGGTACCCAGAGAGCGGCTACTGACCCAC 454
Qy 41 LeuLeuAspHisTyrrAsnValLeuGluArgProValValAsnGluSerAspProLeuGln 60
Db 455 CTATTGACCACTACAAACGCTACTGGAGAGCGCCGCTGCTCAACGAGACGACCGCTG 514
Qy 61 LeuSerPheGlyLeuThrLeuMetGlnIleAspValAspValAspGluLysAsnGlnLeu 80

```

```

Db 515 CTCCTCTTCGCGCTCACCTCATGCAGATCATCGAGTGGAGAGAACACAGCTTTTA 574
Qy 81 IleThrAsnIleTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSer 100
Db 575 ATAACAACATCTGGCTAAACTAGAGTGAATGATATGAACCTTGAGGTGGACACTTCA 634
Qy 101 AspPheGlyGlyValLysAspLeuArgValProProHisArgLeuTrpLysProAspVal 120
Db 635 GATTTCGCGCGGGTCAAGATTTAAGAGTTCACCCACACAGACTATGGAACACGACGTC 694
Qy 121 LeuMetTyrrAsnSerAlaAspGluGlyPheAspSerThrTyrrProThrAsnValVal 140
Db 695 CTTATGTACAACACGCGGAGCGAAGGTTTCAGACGACGATTCACACGAGCTGGTGTG 754
Qy 141 ArgAsnAsnGlySerCysLeuTyrrValProProGlyIlePheLysSerThrCysLysIle 160
Db 755 CGGAACAACGGCTCGTGTCTGTACGTGCGCGCGGCATCTTCAAGAGCACCTGCAAGATC 814
Qy 161 AspIleThrTrpPheProPheAspAspGlnArgCysGluMetLysPheGlySerTrpThr 180
Db 815 GACATCACCTGGTTCCCTTCGACGACCAACGATGCGAGATGAAGTTTGGCAGCTGGACT 874
Qy 181 TyrrAspGlyTyrrGlnLeuAspLeuGlnLeuAspGluGlyGlyGlyAspIleSerSer 200
Db 875 TATGATGGTTATCATGTTGGATCTACACTACAGATGAGGGGCGGAGATATAGCACT 934
Qy 201 PheValThrAsnGlyGluTrpGluLeuIleGlyValProGlyLysArgAsnGluIleTyrr 220
Db 935 TTTGTACGATGCGGAATGGGAGTTAATAGAGTCCCGCAAGCGCACGAGATCTAC 994
Qy 221 TyrrAsnCysCysProGluProTyrrIleAspIleThrPheAlaValIleArgArgLys 240
Db 995 TACAACCTGTTGTCGAGAGCCATACATCGACATCAGTTTTCGGTGGTGTACCGAGGAAA 1054
Qy 241 ThrLeuTyrrTyrrPhePheAsnLeuIleValProCysValLeuIleAlaSerMetAlaLeu 260
Db 1055 ACGCTCTACTACTTCTTCAATCTGATCGTGCCTGCGTGTCTCATCGCTCCATGCTCTA 1114
Qy 261 LeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeuGlyValThrIleLeu 280
Db 1115 TTGGGGTTTCACTTCCCTCCAGACTCCGAGAGAAAAGTTGTCTTAGGTGTGACGATATTA 1174
Qy 281 LeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAla 300
Db 1175 CTGTCTGTGACGGTGTCTTCAATGTCGCGGAGACGATGCCAGGACGCTCGAGCGCC 1234
Qy 301 ValProLeuLeuGlyThrTyrrPheAsnCysIleMetPheMetValAlaSerSerValVal 320
Db 1235 GTGCCCTTGTCTGGCACCTTCACTGCATCATGTTTCATGGTGGCTTCTCTCGTGTCT 1294
Qy 321 SerThrIleLeuLeuLeuAsnTyrrHisArgHisAlaAspThrHisGluMetSerAsp 340
Db 1295 TCCACCATATCTGATCTCAACTACCAACCGGCGACGACACTACGAAATGAGTGTAT 1354
Qy 341 TrpIleArgCysValPheLeuTyrrTrpLeuProTrpValLeuArgMetSerArgProGly 360
Db 1355 TGGATTCTGTGCTGTTCTTTTATGCTGCTGGGTGTGTCGATGTCAGCGCCCGCC 1414
Qy 361 SerAlaThrThrProProAlaArgValProProProAspLeuGluLeuArgGlu 380
Db 1415 TCGCGACGACGCGCGCGCGCGGTACTTCCGCGCGCGACCTCGGAGCTGGAGCTGCGCG 1474
Qy 381 ArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspAspPheArgHisPro 400
Db 1475 CGTCTCTCCAAAGTCGTCTTAGGAAACGTCTCGACATCGATGACGACTTCCGCGCC 1534
Qy 401 GlnAlaGlnGlnProGlnCysCysArgTyrrTyrrArgGlyGlyGluGluAsnGlyValArg 420
Db 1535 CAAGCGCAGCAGCGCGCAATGCTCCGATATACTACAGGGGGGTGAGGAGATGCGCGGG 1594
Qy 421 LeuAlaAlaHisSerCysPheGlyValAspTyrrGluLeuSerLeuLeuLeuLeuLeu 440

```

```

Db      1595 TTGGCGCGCACAGTTCCTCGGTGTCGACTACGAGCTCCCTCATCTTCAAGGAGATT 1654
Qy      441 ArgValIleThrAspGlnMetArgLysAspAspGluAspAlaAspIleSerArgAspTyr 460
Db      1655 AGAGTCATCATCAGATCGCGCAGGACGACGACGATCGGACATTTCCGCGCAGCTGG 1714
Qy      461 LysPheAlaAlaValValAspArgLeuCysLeuIleIlePheThrLeuPheThrIle 480
Db      1715 AAGTTCCGCCGCGCARGGCGGACAGACTTGCCTTATTATCTTTACCTGTTCAACATC 1774
Qy      481 IleAlaThrLeuAlaValLeuLeuSerAlaProHisIleMetValSer 496
Db      1775 ATGCCACGCTAGCGGTGCTGCTCGCGCCACACATCATGCTGTGCG 1822

RESULT 4
AF272778      2907 bp      mRNA      linear      INV 29-APR-2002
LOCUS      Drosophila melanogaster nicotinic acetylcholine receptor Dalpha5
DEFINITION      subunit (nAcRalpha-34E) mRNA, nAcRalpha-34E-A allele, complete cds.
ACCESSION      AF272778
VERSION      AF272778.1 GI:20152839
KEYWORDS
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM      Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 2907)
AUTHORS      Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.
TITLE      Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a
New and Highly Conserved Target of Adenosine Deaminase Acting on
RNA-Mediated A-to-I Pre-mRNA Editing
Genetics 160 (4), 1519-1533 (2002)
JOURNAL
MEDLINE      21969411
PUBMED      11973307
REFERENCE      2 (bases 1 to 2907)
AUTHORS      Grauso, M. and Sattelle, D.B.
TITLE      Direct Submission
JOURNAL
FEATURES
source      1..2907
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="2L; 34E4-5"
/dev_stage="embryo"
1..2907
/gene="nAcRalpha-34E"
/allele="A"
289..2712
/gene="nAcRalpha-34E"
/notes="ion channel; neurotransmitter transmembrane
receptor"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalpha5
subunit"
/protein_id="AAMI3390.1"
/db_xref="GI:20152840"
/translation="MKNALKLTVEVDDDELAVRLACSSNFSSTSTSTSSNORH
NQQLTIQPSRLSTKHHNSIASEQHNSQQQEPASKDEEDVANHGNSNDQQTLLQLOLSS
NMLSPKTAATAAGDEATQQPTNIELCKARKQRRLRRRRKRPATNETDIRKQOOL
SMFPFKTSDTSTPAATSCPTFMQCRASDNFSPISRHDRVSTATTFAWVLIH
VLOVLLVLOQWQVQSVLLFRRIAASATFIAPISVLGFAAOLKNSSSSSSSNSS
NNSQTLLGNLKHWSIFLLIYLNLSAKVCLAGVHEKRLHLLDLPNTLERPVINES
DPLQSLFGLTLMQIDVDENQLLVNTVMALKLEWDMLEWNTSDYGVKDLRIPPHR
IWRDVLNMSADEGPGTQTQNVVVRNNSCLVPPGIFKSTKIDITWFPDDQRC
EMKPGSYDGFQDLQLOQDETGISYVLNGEWELLVPGKRNELIYVNCPEPEYID
ITFAIIRRLTYFFNLIIPCVLIASMALIGFTLPDPSGKLSLGYTILLSLTVLIN
MVAETMPATSDAVPLGTYFNCIMFVASSVSVTLINLVHNRNADTHEMSEWIRVF
LCWLPWILRMSRPGRLILEFPPTTPCSDTSSERKHQILSDVELKERSKSLIANVLID

```

variation

variation

variation

variation

variation

variation

variation

ORIGIN

Alignment Scores:

```

Pred. No.:      4.47e-189      Length:      2907
Score:          2064.50      Matches:      396
Percent Similarity: 82.95%      Conservative: 32
Best Local Similarity: 76.74%      Mismatches: 49
Query Match:     77.47%      Indels:      39
DB:              3          Gaps:      6

```

US-09-303-232-4 (1-496) x AF272778 (1-2907)

```

Qy      17 LeuLeuLeuLeuCysLeuLeuTyrProArgGlyAlaArgCysGlyTyrHisGluLysArg 36
Db      1171 TTATTGATATATTGAATTTATCTGCTAAAGTTTGGCTAGCAGGATATCATGAAAAGAGA 1230
Qy      37 LeuLeuHisHisLeuLeuAspHisTyrAsnValLeuGluArgProValValAsnGluSer 56
Db      1231 CTGTTACACCATCTTTTGGATCCTTATAATACACTAGAACGTCCTCCGTCTCAATGATCG 1290
Qy      57 AspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleIleAspValAspGluLys 76
Db      1291 GACCCGTTACAATTAAGCTTTGGTTTAACTTTAATCAATTAATCGAATATCGATGTGGACGAGAA 1350
Qy      77 AsnGlnLeuLeuIleThrAsnIleTyrLeuLysLeuGluTyrAsnAspMetAsnLeuArg 96
Db      1351 AATCAATTTGATGTCATTAATGTGGTTAAACTGAGTGGAAACGACATGAATCTCCGC 1410
Qy      97 TrpAsnThrSerAspPheGlyGlyValLysAspLeuArgValProProHisArgLeuTyr 116
Db      1411 TGGAAACCTCCGACTATGCGGAGTTAAGCATCTCGAATAACGCGCATCGCATCTGG 1470
Qy      117 LysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspSerThrTyrProThr 136
Db      1471 AAGCCGACGCTGCTGATGTACAACAGTCGCGATGAGGGATTTGACGGCACCTACCACAGC 1530

```

```

DDDFRNCBMPTEGGTLPHPNPAFYRTVYGGDDGSGIGTIGSTRMPDAVTHHTCIKSSST
EYELGULKEIRFITDQRKDDCNDIANDWKFAAMVVDRLCLLIITFMFAILLATIAVL
LSAPHHIVS"
373
/genes="nAcRalpha-34E"
/notes="compared to B allele"
/replace="a"
375
/genes="nAcRalpha-34E"
/notes="results in phenylalanine to isoleucine
substitution; compared to B allele"
/replace="c"
827
/genes="nAcRalpha-34E"
/notes="results in threonine to isoleucine substitution;
compared to B allele"
/replace="t"
1079..1080
/genes="nAcRalpha-34E"
/notes="results in lysine to arginine substitution;
compared to B allele"
/replace="gg"
1081..1086
/genes="nAcRalpha-34E"
/notes="results in deletion of asparagine and serine;
compared to B allele"
/replace="n"
1119..1120
/genes="nAcRalpha-34E"
/notes="results in serine, glycine, and asparagine
insertion; compared to B allele"
/replace="agcggcaac"
1121
/genes="nAcRalpha-34E"
/notes="results in asparagine to serine substitution;
compared to B allele"
/replace="g"

```

```
Qy 137 AsnValValValArgAsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSer 156
Db 1531 AACGTGGTGGTGGCGGAACAACGGCTCGTGTCTATACGTTCCGCCGGGATCTTCAGTCG 1590
Qy 157 ThrCysLysIleAspIleThrTrpPheProPheAspGlnArgCysGluMetLysPhe 176
Db 1591 ACGTGCAGATCGACATCACGTGGTTCCTTCGATGACACGCGGTGGAGATGAAGTTC 1650
Qy 177 GlySerThrThrTyrAspGlyTyrGlnLeuAspLeuGlnLeuGlnAspGluGlyGly 196
Db 1651 GCGAGTTGAGACTGACGCGGATTCAGCTGGATTTACAAATTACAGATGAACACGTCGGGT 1710
Qy 197 AspIleSerSerPheValThrAsnGlyGluTrpGluLeuIleGlyValProGlyLysArg 216
Db 1711 GATATCAGCAGTACGTGCTCAACGCGGAGTGGGAACCTACTGGGTGTCGCCGCAACAGT 1770
Qy 217 AsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPheAlaVal 236
Db 1771 AACGAGATCTATTACAACCTGCTGCCGGAACCTATATAGACATCACCTTCGCCATCATC 1830
Qy 237 IleArgArgLysThrLeuTyrTyrPhePheAsnLeuIleValProCysValLeuIleAla 256
Db 1831 ATCCGCCGACGAAACACTGTACTATTCTTCAACCTGATCATACCTTGTGTACTGATGCC 1890
Qy 257 SerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeuGly 276
Db 1891 TCCATGGCTTGTCTCGGATTCCACCTGCCGCGCAGATTCGGGTGAAAAAATTATCGCTGGT 1950
Qy 277 ValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAla 296
Db 1951 GTTACCATCTGTCTCTCGCTCAGCGGTGTTCTCGAATATGTTGGTGGCAGACAATCCGGCT 2010
Qy 297 ThrSerAspAlaValProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAla 316
Db 2011 ACTTCCGATGCGGGGCCATTGCTGGGTACATATTTCAATTGCATATGTTTATGTAGCT 2070
Qy 317 SerSerValValSerThrIleLeuLeuAsnTyrHisHisArgHisAlaAspThrHis 336
Db 2071 TCATCCGTTGTGTCAACGATTTAATAATTAATATCATCATCGAATGCTGATACGCAC 2130
Qy 337 GluMetSerAspTrpIleArgCysValPheLeuTyrTrpLeuProTrpValLeuArgMet 356
Db 2131 GAAATGTCGAATGGAACGATCGTGTGTTTGTGCTGGCTGCCATGATATTCGAATG 2190
Qy 357 SerArgProGlySerAlaThrProProProAlaArgValProProProProProPro 374
Db 2191 AGTCGCCGAGA-----CGACCGCTGATCTCTAGAGTCCGACACACCCCTGTCTG 2241
Qy 375 -----ProGlnAlaGlnGlnProGlnCysCysArgTyrTyrArg----- 412
Db 2242 GACACATCTCCGAGCGGAAGCACACGATCTCTCCGACGTTGAGCTGAAGAGCGCTCG 2301
Qy 383 SerLysSerLeuLeuAlaAsnValLeuAspIleAspAspAspPheArgHis----- 399
Db 2302 TCGAATCGCTGCTGGCCAGCTACTAGACATCATCATGATGATCTCCGSCACAATGTGCG 2361
Qy 400 -----ProGlnAlaGlnGlnProGlnCysCysArgTyrTyrArg----- 412
Db 2362 CCCATGACGCCCGCGGAACACTGCCACACACCCCGGCTTCTATCGACGGTTATGGA 2421
Qy 413 GlyGlyGluGluAsnGlyAlaGly-----LeuAla 422
Db 2422 CAAGGCGACGATGGCAGCATTTGGGCCCAATTGGCAGCACCCCAATGCGGATCGCGTCA 2481
Qy 423 AlaHisSerCysPhe-----GlyValAspTyrGluLeuSerLeuIleLeuLysGluIle 440
Db 2482 CATCATAGTGCATCAATCATCATCACTGAATATGAATAGTTTAACTTTAAAGAAATT 2541
Qy 441 ArgValIleThrAspGlnMetArgLysAspAspGluAspAlaAspIleSerArgAspTrp 460
Db 2542 CGCTTTATACTGATCAGCTAGCTAGTAAGATGACGAGTGCAATGACATTTGCCAATGATGG 2601
Qy 461 LysPheAlaAlaMetValValAspArgLeuLeuLeuIlePheThrLeuPheThrIle 480
```

```
Db 2602 AAATTGACGATGTCGTTGACAGACTGTCCTTATCATATTCACATGTCGCATA 2661
Qy 481 IleAlaThrLeuAlaValLeuLeuSerAlaProHisIleMetValSer 496
Db 2662 TTAGCCACAATAGCTGTACTACTATCATCAGCACCATATTAATGTCCTG 2709

RESULT 5
DME554210 1683 bp mRNA linear INV 01-APR-2003
LOCUS Drosophila melanogaster mRNA for nicotinic acetylcholine receptor
DEFINITION subunit Dalp7 (nAcRalpha-18C gene).
ACCESSION AJ554210
VERSION AJ554210.1 GI:29466436
KEYWORDS nAcRalpha-18C gene; nicotinic acetylcholine receptor subunit
Dalp7.

SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1683)
REFERENCE Millar,N.S.
AUTHORS Direct Submission
TITLE Submitted (31-MAR-2003) Millar N.S., Department of Pharmacology,
JOURNAL University College London, Gower Street, London, WC1E 6BT, UNITED
KINGDOM

FEATURES
source Location/Qualifiers
1..1683 /organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
1..1683 /gene="nAcRalpha-18C"
26..1654 /gene="nAcRalpha-18C"
/codon_start=1
/product="nicotinic acetylcholine receptor subunit
Dalp7"
/protein_id="CAD86936.1"
/db_xref="GI:29466437"
/translation="MSFPQHSLEPEATANGRMVLYGLLIMIPACAAGPHEKRLRH
ALLDYNLSRPVYVNSDPLQLSGLTMOITDVENKQLITINILKWNMDNLRW
NSSEFGVGRDLRIPPHRLWKPDLVYNASDEGDTATVNVNRNGSLVYVPQIFK
STCKIDITWPFDDORCEMGKGSWTYDGFOLDLOQDEAGDISFTNGSWDILGVP
KXNEIYVNCPPXYIDITFALLRRLKTYFFNLIVPCVLIAKALLGFTLPDPSGE
KLSGLVILSLIVFLNMVAETMPATSDAVPLGKFNCFMFWASSVSTLILVNIYH
HRNPDTHEMSEWIRVFLYWLPCILRMQRFQGVGECPPPPSSSSASGKKQOIQN
VELKERSKSLANLVIDDDPCNHRCSATLPQPTTYRTMYRQDGDGSGVGPQA
GPVVDGRLEHAIHSHTCLTSSAEYELALILKELRWITLQKKEDETSITRDWKEAAMV
VDRLLIIFITLTIATLAVLSAPHFIVSGVRG"
26..131
/gene="nAcRalpha-18C"
132..1651
/gene="nAcRalpha-18C"
/product="nicotinic acetylcholine receptor subunit
Dalp7"

sig_peptide
mat_peptide

ORIGIN
Alignment Scores:
Pred. No.: 1,01e-187 Length: 1683
Score: 2047.00 Matches: 399
Percent Similarity: 81.16% Conservative: 36
Best Local Similarity: 74.44% Mismatches: 47
Query Match: 76.81% Indels: 54
DB: 3 Gaps: 9

US-09-303-232-4 (1-496) x DME554210 (1-1683)
Qy 2 GlyGlyArgAlaArgArgSerHisLeuAlaAlaProAlaGlyLeuLeuLeuLeuCys 21
Db 71 GGTGGCAGATG-----CTGGTCTATGGCTGGGACTTTTAATTATGATA--- 115
```

QY 22 LeuLeuTrpProArgGlyAlaArgCysGlyTyrHisGluLysArgLeuLeuHisHisLeu 41
Db 116 -----CCGGCTTGCGGCTGGACCCCATGAGAGCGGCTACTCCAGCCCTT 163
QY 42 LeuAspHisTyrAsnValLeuGluArgProValValAsnGlnSerAspProLeuGlnLeu 61
Db 164 CTGGACAACATACACAGCGCTCGAGCGTCCGGTGGTCAATGAATCCGATCCATTGCAACTG 223
QY 62 SerPheGlyLeuThrLeuMetGlnIleLeuValAspValAspGluLysAsnGlnLeuLeu 81
Db 224 AGCTTCGGACATAACATCATCGAGATTATCGATGGACGAAAGAAATCAACTGCTTATA 283
QY 82 ThrAsnIleTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAsp 101
Db 284 ACGATATTGGCTTCAAAATGGAATGAACATATGAATCTTCGATGGAATTCAGGTGAG 343
QY 102 PheGlyGlyValLysAspLeuArgValProProHisArgLeuTrpLysProAspValLeu 121
Db 344 TTCGGTGGTGGCGGATCTGCGAATTCGCGCACATCGCTATGGAACCGGATGTACTG 403
QY 122 MetTyrAsnSerAlaAspGluGlyPheAspSerThrTyrProThrAsnValValArg 141
Db 404 ATGTACAAACATGTCGACGAGGGCTTCGATGGAACTGACGCCACAAATGTGGTTCGC 463
QY 142 AsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSerThrCysLysIleAsp 161
Db 464 AATATCGGAGCTGTCTGTACGTACCGCCAGGTATATTTAAGTCAACGTGAATCGAC 523
QY 162 IleThrTrpPheProPheAspAspGlnArgCysGluMetLysPheGlySerThrTyr 181
Db 524 ATTACGTGGTTCATTCGACGATCAGAGATGTGAATGAAATTTGGTTCGTGACCTAC 583
QY 182 AspGlyTyrGlnLeuAspLeuGlnLeuGlnAspGluGlyGlyAspIleSerSerPhe 201
Db 584 GATGGGTTTCAGTTGGACCTGCAGTTGCAGACAAAGCTGTGGCGACATTTCTAGCTTT 643
QY 202 ValThrAsnGlyGluTrpGluLeuIleGlyValProGlyLysArgAsnGluIleTyr 221
Db 644 ATACCAATGGGATGGGACTTGTAGGTGTGCCGGTAAACGAAATGAATCTACTAT 703
QY 222 AsnCysCysProGluProTyrIleAspIleThrPheAlaValValIleArgAtGLysThr 241
Db 704 AATTGCTGCCAGAACCTTATATTGACATAACATTCGCCATTTGTATAAGCGCAAAACG 763
QY 242 LeuTyrTyrPhePheAsnLeuIleValProCysValLeuIleAlaSerMetAlaLeuLeu 261
Db 764 TTGTACTATTTTTCATCTCTGATTGTCGGTGGTACTGATCGCTCCATGGCACTGCTA 823
QY 262 GlyPheThrLeuProProAspSerGlyLysLeuSerLeuGlyValThrIleLeuLeu 281
Db 824 GGGTTTACACTGCCACAGATTCTGGTGAAAGCTTTTCGCTTGAGTTACAATTTCTATA 883
QY 282 SerLeuThrValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaVal 301
Db 884 TCGCTTACAGTCTCTCTCAACATGTCGCGCAACAATCGCGGACCTCCGATCGGTA 943
QY 302 ProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAlaSerSerValSer 321
Db 944 CGCTGTCGGTAAAGTATTTCAATTGCAATTATGTTTATGTTGGCTCATCATGTTGTCA 1003
QY 322 ThrIleLeuIleLeuAsnTyrHisArgHisAlaAspThrHisGluMetSerAspTrp 341
Db 1004 ACCATCTTGTCTCAATTTATCATATGAAATCCAGATACGATGAAATGAGTGAATGG 1063
QY 342 IleArgCysValPheLeuTyrTrpLeuProTrpValLeuArgMetSerArgProGlySer 361
Db 1064 ATAAGAGTAATATCTTTTATTGTTTACCTTTCATATTCGATGTCATGCAAGACCGGACAG 1123
QY 362 AlaThrThrProProAlaArgValProProPro----- 374
Db 1124 GTT-----GGCTACGAATGTCCGCGCGCCCTCTTCTCGAGTTCCTCGCA 1171
QY 375 -----AspLeuGluLeuArgGluArgSerSerLysSer 385

Db 1172 TCCGGCGAGAAGAACAGATCCAAACCGTTCAGCTCAAGGAGAGTCTCCTCAAGTCT 1231
QY 386 LeuLeuAlaAsnValLeuAspIleAspAspPheArg-----HisProGlnAla--- 402
Db 1232 CTGCTGGCCCAATGTGCTCGATATAGACGATGATTTCCGATGCAATCATCGATGTGCCAGC 1291
QY 403 -----GlnGlnProGlnCysCysArg---TyrTyrArgGlyGlyGluGluAsn 417
Db 1292 GCGACTTTGGCCACCACCGCCACATATTACAGGACGATGTACAGGCAAGGGATGACGGC 1351
QY 418 GlyAlaGly-----LeuAla 422
Db 1352 AGCGTGGGACCGGTGGGACCGAGCTGTCAGTGTGGACCGGCGCTTTGACAGAGGCCATT 1411
QY 423 AlaHisSerCysPhe-----GlyValAspTyrGluLeuSerLeuLeuLeuLysGluIle 440
Db 1412 TCCACACACTGTCTGACATCTCTCGGAGTACGAATCTGGCGCTGATACTCAAGAGAGCTG 1471
QY 441 ArgValIleThrAspGlnMetArgLysAspAspGluAspAlaAspIleSerArgAspTrp 460
Db 1472 CGTGGTAATACAGAACAGCTCAAAAAAGAGACGAAACAAGCGACATTACCGAGATTGG 1531
QY 461 LysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrLeuPheThrIle 480
Db 1532 AATTTCTGCCATGGTGTGCGATCGTGTGCTTATTTTACCTTTTACTATT 1591
QY 481 IleAlaThrLeuAlaValLeuLeuSerAlaProHisIleMetValSer 496
Db 1592 ATAGCAACCTCGCTGTACTTCTCTCAGCGCGCATTTTCATTGTGAGT 1639
RESULT 6
AY036613 2834 bp mRNA linear INV 30-APR-2002
LOCUS Drosophila melanogaster nicotinic acetylcholine receptor Dalphas
DEFINITION subunit (nAcRalpha34E) mRNA, nAcRalpha34E-B allele, complete cds,
alternatively spliced.
AY036613
AY036613.1 GI:20340268
Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 2834)
AUTHORS Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.
TITLE Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
Dalphas5, Dalphas6 and Dalphas7, in Drosophila melanogaster Identify a
New and Highly Conserved Target of Adenosine Deaminase Acting on
RNA-Mediated A-to-I Pre-mRNA Editing
JOURNAL Genetics 160 (4), 1519-1533 (2002)
MEDLINE 21969411
PUBMED 11973307
REFERENCE 2 (bases 1 to 2834)
AUTHORS Grauso, M. and Sattelle, D.B.
TITLE Direct Submission
JOURNAL Submitted (26-MAY-2001) MRC-FGU Human Anatomy and Genetics,
University of Oxford, South Parks Road, Oxford OX1 3QX, UK
FEATURES
source Location/Qualifiers
1..2834
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="34E4-34E5"
/dev_stage="embryo"
1..2834
/gene="nAcRalpha34E"
/allele="B"
289..1737
/gene="nAcRalpha34E"
/note="ion channel; neurotransmitter transmembrane

receptor; exon 5 is excluded due to exon skipping, which generates a loss of reading frame and a truncated polypeptide; alternatively spliced"

```

/codon_start=1
/product="nicotinic acetylcholine receptor Dalpha5
subunit"
/protein_id="AAK67256.1"
/db_xref="GI:203403269"
/translations="KKNHAKLITEVDDEDLWLAHLCHCSNISSSSSTRTTSSNORH
NOQLTLPQSLSTGKXSHNIAESKHNSSOOEPAKDEAVNHNKRSNQOHTLOGLDS
NMLSEKTAATAAAGDEATQPTNIRLCARQRLRRRRKCAPETNEDTKKQQL
SMPEKTKSTKSTGYTPAALTSCPTNATVQCASDNESIPLSRHDSVSTATFAMWL
VLQVLVLSQWLQWLVHQKRSVLFRNIAASTNYSFGLVHKRLHLDDLPNTLRVLNE
SSQSTQITGLNKLHNSWIFLLYINLSAKVCLUAGHVKRLHLDDLPNTLRVLNE
SDPLQSFGLTMOIIVDKQKQLVNLVWLKXVHNRLLHDDLPNTLRVLNEIPH
RKWPKLVYDNGSADEGDTGYTNVWRNGLSPVPGIKFSTCKIDITWFFPDDOR
CEMKGFSWTVDFGQVCPANVTSTITTAARN?"

```

ORIGIN

Alignment Scores:		
Pred. No.:	2.35e-174	2834
Score:	1942.00	374
Percent Similarity:	78.29%	Conservative: 30
Best Local Similarity:	72.48%	Mismatches: 48
Query Match:	71.74%	Indels: 65
DB:	3	Gaps: 7
US-09-303-232-4	(1-496)	X AY036613 (1-2834)

Qy	217	Asn	Glul	Ile	Tyr	Tyr	Asn	Cys	Cys	Pro	Glup	Pro	Tyr	Ile	Asp	Ile	Thr	Phe	Ala	Val	Val	236
Db	1854	AAC	GAG	ATC	TAT	TAC	AACT	GTG	CCG	GAA	CCCT	TAT	TAG	ACAT	CAC	CTT	CGC	ATC	ATC	ATC	ATC	191
Qy	237	Ile	Arg	Gly	Ser	Thr	Leu	Tyr	Tyr	Phe	Phe	Asn	Ile	Val	Pro	Cys	Val	Leu	Ile	Ala	256	
Db	1914	ATC	CGC	CGC	GCA	AC	CAC	TGT	ACT	TAT	TCT	TCA	CACT	GTAT	CAT	ACC	TGT	TGT	ACT	GAT	TGCC	197
Qy	257	Ser	Met	Ala	Leu	Leu	Gly	Phe	Thr	Leu	Pro	Pro	Asp	Ser	Gly	Glu	Leu	Ser	Leu	Gly	276	
Db	1974	TCC	ATG	CCCT	TGCT	CGG	ATT	CAC	CCCT	CGC	CGC	GAT	TCC	GGGT	GAAA	AAAT	TAT	TCG	CGT	GGGT	203	
Qy	277	Val	Thr	Ile	Leu	Leu	Ser	Leu	Thr	Val	Phe	Leu	Asn	Met	Val	Ala	Glut	Thr	Met	Pro	Ala	296
Db	2034	GTT	ACC	ATC	TTC	TCT	CGCT	CAC	CGT	GTT	TCT	GAT	TAT	GTT	GTT	CGC	GAG	ACA	TGCG	CGGT	209	
Qy	297	Thr	Ser	Asp	Ala	Val	Pro	Leu	Leu	Gly	Thr	Tyr	Phe	Asn	Cys	Ile	Met	Phe	Met	Val	Ala	316
Db	2094	ACT	TCC	GAT	CGGT	GCC	ATTG	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	211	
Qy	317	Ser	Ser	Val	Val	Ser	Thr	Ile	Leu	Ile	Leu	Asn	Tyr	His	His	Arg	His	Ala	Asp	Thr	His	336

317	QY	SerSerValValSerThrIleLeuLeuAsnTyrHisHisArgHisAlaAspThrHis	333
2114	Db	-----	2114
337	QY	GluMetSerAspTrpIleArgCysValPheLeuTyrTrpLeuProTyrValLeuArgMet	356
2115	Db	-----TGGATACGCATCGTCTTTTGTGCTGGCTGCCATGGATATTCGAATG	2165
357	QY	SerArgProGlySerAlaThrThrProProAlaArgValProProProPro-----	374
2163	Db	AGTCGCCACGGA-----CGACCGCTGATCCTAGAGTTCCCGACCAAGCCCTGTTTCG	2213

2163	Db	AGTCGCCAGGA-----CGACCGCTGATCCTAGAGTTCCCGACACACCGCCCTGTTTCG	2213
375	QY	-----AspLeuGluLeuArgGluArgSer	382
2214	Db	GACACATCTCCGAGCGGAAGCACACAGATACTCTCCGACGTTAGCTGAAGAGCGCTCG	2277
383	QY	SerLysSerLeuLeuAlaAsnValLeuAspPileAspAspPheArgHis-----	399
2274	Db	TCGAAATCGCTGCTGGCCAAAGTACTAGACATCGATGATGACTTCCGGCACAAATTGTGCG	2333
400	QY	-----ProGlnAlaGlnProGlnCysCysArgTyrTyrArg-----	412
2334	Db	CCCATGAGCGCCGGGGAACACTGGCCACACACCCCGCTTTCTATCGACGGTTTATGGA	2399
413	QY	GlyGlyGluGluAsnGlyAlaGly-----LeuAla	422
2394	Db	CAAGGCGCAGATGGCAGCATTTGGGCCAATTGGCAGACCCCGAATGCCGATGCCGTACC	2455
423	QY	AlaHisSerCysPhe-----GlyValAspTyrGluLeuSerLeuLeuLeuLysGluIle	440
2454	Db	CATCATCGTCATCAACATCATCAACTGAATATGAATTAGTGTAAATCTTTAAAGAAAT	2513
441	QY	ArgValIleThrAspGlnMetArgLysAspAspGluAspAlaAspIleSerArgAspTyr	460
2514	Db	CGCTTTATACTGATCAGCTACGTAAAGATGACGAGTGCATGCATTCGCAATGATTGG	2573
461	QY	LysPheAlaAlaMetValValaspArgLeuCysLeuIleIlePheThrLeuPheThrIle	480
2574	Db	AAATTTGCAGCTATGTCGTTGACAGACTGCTGCTTATCATATTATTCACAAATGTCGCAATA	2633
481	QY	IleAlaThrLeuAlaValLeuSerAlaProHisIleMetValSer	496
2634	Db	TTAGCCAAATAGCTGTACTACTATCGGCACCATATTATTGTTCTCG	2681
RESULT 8			
AX009610			
LOCUS			
DEFINITION			
Sequence 1 from Patent EP0962528.			
ACCESSION			
AX009610			
VERSION			
AX009610.1 GI:9996842			
KEYWORDS			
Drosophila melanogaster (fruit fly)			
SOURCE			
linear			
PAT 06-SEP-2000			

RESULT 8					
AX009610					
LOCUS					
DEFINITION	AX009610	2886 bp	DNA		
ACCESSION	Sequence 1 from Patent EP0962528.				
VERSION	AX009610.1				
KEYWORDS	GI:9996842				
SOURCE	.				
ORGANISM	Drosophila melanogaster (fruit fly)				
	Drosophila melanogaster				
	linear				PAT 06-SEP-2000


```
Db 1175 ACAGGATGAGGAGCTGGAAGTCTGAGAGAGTCTGCGAAGTCTTCTGCGCGAATGTT 1234
Qy 391 LeuAspIleAspAspPheArgHis-----ProGlnAlaGlnGln----- 404
Db 1235 CTAGATATTGATGATGACTTCAGACACAGCGCCCTCCGCTCTTAACAGTACTGCTCGACC 1294
Qy 405 -----ProGlnCysCysArgTyr-----TyrArgGlyGly----- 414
Db 1295 GGGAAATTGGACCTGGGTGCTCAATATTCGCGACGGATTTCCTCGGTGCTGTCGCGT 1354
Qy 415 -----GluGluAsnGlyValaGlyLeuAlaAlaHisSerCysPheGlyValaAsp 430
Db 1355 CCGTCCACGATGGAAGACGTGGCGGGCTGGGTAGCCAC-----CAT 1399
Qy 431 TyrGluLeuSerLeuIleLeuLysGluIleArgValIleThrAspGlnMetArgLysAsp 450
Db 1400 CGCGAGCTGCACCTCATATGAGAGAGCTGCAGTTCATCAGCGCAGGATGAAGAGGCT 1459
Qy 451 AspGluAspAlaAspIleSerArgAspTrpLysPheAlaAlaMetValValaAspArgLeu 470
Db 1460 GATGAGGAGCCGAGCTGATCAGCAGCTGGAGATTGCTGGGATGGTTGTTGATAGGTTT 1519
Qy 471 CysLeuIleIlePheThrLeuPheThrIleIleAlaThrLeuAlaValLeuSerAla 490
Db 1520 TGCCTGTTGCGTTCACACTTTTCACAATCATCGGACAGTAGTGTCTCTGTTATCGGCA 1579
Qy 491 ProHisIleMetVal 495
Db 1580 CCGCATATCATCGTG 1594
```

```
RESULT 11
AX009614
LOCUS AX009614 3109 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 5 from Patent EP0962528.
ACCESSION AX009614
VERSION AX009614.1 GI:9996846
KEYWORDS
SOURCE
ORGANISM Heliothis virescens (tobacco budworm)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE 1
Adamczewski M.D., Schulte J.D. and Oellers, N.D.
Nucleic acids encoding acetylcholin-receptor subunits from insects
Patent: EP 0962528-A 5 08-DEC-1999;
BAYER AG (DE)
```

```
FEATURES
source
1..3109
location/Qualifiers
1..3109
/organism="Heliothis virescens"
/mol_type="unassigned DNA"
/db_xref="taxon:7102"
95..1600
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAC07501.1"
/db_xref="GI:9996847"
/db_xref="REFMREMBL:CAC07501"
/translation="MAPLAAALALLPVSEQGPHEKRLINALLANNTLPRPVANE
SELEVRFLGTLQIIDVDEKNQILLITNIWLSLEWYDNLNRWNSDEYGVKDLRTPN
KLKPDVLYNSADEGPDGTQTNVNVRSGLYVPPGIPKSTCKMDIAWFPDDOH
CDMKFGSWTYNGQLDLVADAGGLSDPITNGEWYLVIGMPKKNITVYACPEPYV
DVFTIMRRRLTYFNNLIVPCVLISSMALLGFTLPDPSGEKLTIGVTLISLTIVFL
NLVAETLPQVSDAIPLIGTYFNCIMFWASSVLTVVLYNHHRTADIEHMPQWIKSV
FLOWPLWLRMSRPGKKITRNTIMNTRMRELEKRSKSLANLVIDIDDDFRHGPP
PPNSTAGNLGPGCSFTFRDFFRSFVRPSFMEDVGGGLGSHRELHLILRELQFITA
RMKKADEAEELISDMKFAAMVWDRFCLFVTLFIATVALVLSAPHIIIVQ"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 2,4e-164 Length: 3109
Score: 1809.00 Matches: 352
```

```
Percent Similarity: 77.52% Conservative: 55
Best Local Similarity: 67.05% Mismatches: 78
Query Match: 67.88% Indels: 40
DB: 6 Gaps: 8
US-09-303-232-4 (1-496) x AX009614 (1-3109)
```

```
Qy 5 AlaArgArgSerHisLeuAlaAlaProAlaGlyLeuLeu----- 17
Db 38 GCGCGCCGCGGTGCTAGCGAGCGTGGCGGGTGTGGTATGTCGCGCGCGTCCGCGCTATG 97
Qy 18 -----LeuLeuLeuCysLeuLeuTrpProArgGlyAlaArgCysGly 31
Db 98 GCCCCTATGTTGGCGGCTTGGCGCTGCTGGCTTGTGCCCTATCGAGCAA---GGT 154
Qy 32 TyrHisGluLysArgLeuLeuHisLeuLeuAspHisTyrAsnValLeuGluArgPro 51
Db 155 CCTCAGAGAAGAGACTCTTGAACGCGTGTCTGCGGAACACACACCTTGAGGCGGACCG 214
Qy 52 ValValAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleIle 71
Db 215 GTGCCCAACCGAGAGCGAACCCGCTAGAGGTCAGGTTGCGCTTGACCTTGCAACAATCAT 274
Qy 72 AspValAspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeuLysLeuGluTrpAsn 91
Db 275 GACGTGACGAGAGAAGAACTACTATTATACCAANTATATGGCTGTCTGGAGTGGAAAT 334
Qy 92 AspMetAsnLeuArgTrpAsnThrSerAspPheGlyGlyValLysAspLeuArgValPro 111
Db 335 GACTACAACTGAGGTGGAAACGACGAGGATATGGCGGGGTCAAGGACCTCAGGATCAGC 394
Qy 112 ProHisArgLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp 131
Db 395 CCCAACAAAGTTGTGGAAAGCGGACGCTCTTATGTATTAATAGTGTGACGAGGGTTTGAC 454
Qy 132 SerThrTyrProThrAsnValValArgAsnGlySerCysLeuTyrValProPro 151
Db 455 GGGACCTACACAGCAACACGTTGGTGTGCAAGAGCGGCGGAGTGCCTGTAGTGCACCT 514
Qy 152 GlyIlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArg 171
Db 515 GGCATATTCAAGAGACATCAAGATGACATCGCGTGGTTTTCCTTCGACGACCAACAC 574
Qy 172 CysGluMetLysPheGlySerTrpThrTyrAspGlyTyrGlnLeuAspLeuGlnLeuGln 191
Db 575 TGTGATATGAAGTTCGGTAGCTGACATATGACGGCAATCAGTTGGATCTCTGGTCTAAAA 634
Qy 192 AspGluGlyGlyGlyAspIleSerSerPheValThrAsnGlyGluTrpGluLeuIleGly 211
Db 635 GATGAGCGCGCGGCGATCTATCGGACTTCATACAAATGGGAGTGGTATCTAATAGGA 694
Qy 212 ValProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIle 231
Db 695 ATGCCAGGCAAAAAGAACACATAACATACGCGTGTGCTGCCCGGAGCCCTACGTGAGCGTC 754
Qy 232 ThrPheAlaValIleAlaArgArgLysThrLeuTyrTyrPhePheAsnLeuIleValPro 251
Db 755 ACCTTCACCATCATGATGAAGAAGCAACCTTGTACTACTTCTTCAACCTGATCGTCCCG 814
Qy 252 CysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlu 271
Db 815 TGGGTGCTGATCTCATCGATGGCACTCTCTCGGCTTCACATGCCACCACTCCGGAGAG 874
Qy 272 LysLeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAla 291
Db 875 AAACCTCACACTGGAGTCACTATTCTTCTATCGCTGACGGTGTTCCTCAACCTGGTAGCC 934
Qy 292 GluThrMetProAlaThrSerAspAlaValProLeuLeuGlyThrTyrPheAsnCysIle 311
Db 935 GAGACCCCTGCCACAGGTCTCCGACGCTATCCCGCTGTAGGACGCTACTTCAATTGCTATC 994
Qy 312 MetPheMetValAlaSerSerValValSerThrIleLeuIleLeuAsnTyrHisHisArg 331
```


445 GAAAGCTGTCAAGGACCTCATGAAAAAGCGCTGCTGAACCATCTGCTGTCCACCTACAAT 504
47 VallLeuGluArgProValValAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThr 66
505 ACGTGTGAGGACCGGTGGCCAAATGAATCGAGAGCCCTCGAGGTTAAAGTTTCGACTGACG 564
67 LeuMetGlnIleIleAspValAspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeu 86
565 CTGCAGCAGATCATCGATGGAGAGAGAAATCAGCTTCTCAACGAATCTTTGGCTT 624
87 LysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspPheGlyGlyValLys 106
625 TCGTTGGAGTGGAAAGCACTACAATCTCGCTGGAATGAAACGGAATCGGCGGGTCAAG 684
107 AspLeuArgValProProHisArgLeuTrpLysProAspValLeuMetTyrAsnSerAla 126
685 GATCTACGAATCAAGCCCAACAGCTGTGGAAAGCCGACGCTCATGTATCAACAGCGCG 744
127 AspGluGlyPheAspSerThrTyrProThrAsnValValArgAsnAsnGlySerCys 146
745 GATGAGGATTCGATGCGCATCACACCAACATTTGGTCAAAATGGCGGCGATTTGT 804
147 LeuTyrValProProGlyIlePheLysSerThrCysLysIleAspIleThrTrpPhePro 166
805 CTGTACGTGCCCCCTGCTATCTTCAAGAGCACATGCAAGATGGACATCACGTGGTCCCA 864
167 PheAspGlnArgCysGluMetLysPheGlySerTrpThrTyrAspGlyTyrGlnLeu 186
865 TTTGATACCAACAATTCGAAATGAATTCGGTAGTTGGACTTACGATGGAATTCAGTTG 924
187 AspLeuGlnLeuGlnAspGluGlyGlyAspIleSerSerPheValThrAsnGlyGlu 206
925 GATTGGTTTGAATTCGAGATGGAGGGATCTTCCGATTTCAACAAATGGCGGAG 984
207 TrpGluLeuIleGlyValProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGlu 226
985 TGGTACTTGTGTCATGCGCGGGAAGAAGAAATACGATAGTACTACGCTGCTGCCAGAA 1044
227 ProTyrIleAspIleThrPheAlaValIleArgArgLysThrLeuTyrTyrPhePhe 246
1045 CCATATGTCATATCACTTTACTATACAAATTCGTCGCGTACATATATATTTTTC 1104
247 AsnLeuIleValProCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuPro 266
1105 AATTATTCGTGCATGTGCTAATCTCATCGATGCCCTACTGGGCTTCACATGCCG 1164
267 ProAspSerGlyGluLysLeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPhe 286
1165 CCGGATTCGGCGGAGAACTGACGCTGGCGGTAACATCTACTATCATTAACAGTATT 1224
287 LeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValProLeuLeuGlyThr 306
1225 CTAAACCTTGTGCGGAGTCCATGCCGACACGTCGGATGCTGTCTCTATAGGCACC 1284
307 TyrPheAsnCysIleMetPheMetValAlaSerSerValValSerThrIleLeuIleLeu 326
1285 TACTTCAATTCATCATGTTCTGTCCTCGCTCGGTGCTGCAGCAGTAGTGCTC 1344
327 AsnTyrHisHisArgHisAlaAspThrHisGluMetSerAspTrpIleArgCysValPhe 346
1345 AACTACCACCATCGCAGCGAGCATTCACGAGATGCCCGGTGATCAAGTCCGTTTTC 1404
347 LeuTyrTrpLeuProTrpValLeuArgMetSerArgProGlySerAlaThrTrpProPro 366
1405 CTCAATGGCTGCGCTGGATCTTTGCGAATGGGTGCGACCGCGTCGCAAGATTACACGCAA 1464
367 ProAlaArgValPro---ProProAspLeuGluLeuArgGluArgSerSerLysSer 385
1465 ACAATACTATTAGCAATCCGATGAGAGCTGAGCTAAAGAGCGCTCCTCCAAATCC 1524
386 LeuLeuAlaAsnValLeuAspIleAspAspPheArgHisProGlnAlaGlnGlnPro 405
1525 CTGCTGGCCCAATGCTCTCGACATCGACGACGCTTCGCGGCACCAATATCT----- 1575

Qy 406 GlnCysCysArgTyrTyrArgGlyGlyGluAsnGlyAlaGlyLeuAlaAala----- 423
Db 1576 -----GGCTCCCAAAACCGCAATGGCTGCTCGCGCAGCTTC 1611
Qy 424 -----HisSerCysPheGlyValAspTyr---GluLeu 433
Db 1612 GGTGCGGCCCAACGCTGGAGGAGCATCACAGGCCCATCGGCTCAATCAAAAGATCTT 1671
Qy 434 SerLeuIleLeuLysGluIleArgValIleThrAspGlnMetArgLysAspAspGluAsp 453
Db 1672 CATCTAATTCCTCAAGAAATGCAATTTATTACGCGCGGATGCGCAAGCTGACGACGAA 1731
Qy 454 AlaAspIleSerArgAspTrpLysPheAlaAlaMetValValAspArgLeuCysLeuIle 473
Db 1732 GCGCAATGTATCGCGCATTCGGAAGTTCGCGCAATGTTGTGGATAGATTGTTTAAT 1791
Qy 474 IlePheThrLeuPheThrIleIleAlaThrLeuAlaValLeuLeuSerAlaProHisIle 493
Db 1792 GTTTTCAGCTCTTCAGATTATTGCAACGGTTACGGTGCTGCTCTCCGCTCCGCACATA 1851
Qy 494 MetVal 495
Db 1852 ATCGTG 1857
RESULT 13
AF321447
LOCUS
DEFINITION
Drosophila melanogaster nicotinic acetylcholine receptor Dalpha6
subunit variant type III (nAcRalpha-30D) mRNA, complete cds,
alternatively spliced.
AF321447
AF321447.1 GI:20152848
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 2023)
Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.
Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
Dalphas, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a
New and Highly Conserved Target of Adenosine Deaminase Acting on
RNA-Mediated A-to-I Pre-mRNA Editing
Genetics 160 (4), 1519-1533 (2002)
21969411
PUBMED
11973307
2 (bases 1 to 2023)
Grauso, M. and Sattelle, D.B.
Direct Submission
Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional
Genetics Unit, University of Oxford, South Park Road, Oxford OX1
3QX, UK
Location/Qualifiers
1. .2023
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="30D1"
/dev_stage="embryo"
1. .2023
/gene="nAcRalpha-30D"
379. .1863
/gene="nAcRalpha-30D"
/note="ion channel; neurotransmitter transmembrane
receptor; alternatively spliced; contains exons 3b and 8b"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalpha6 subunit
variant type III"
/protein_id="AA13394.1"
/db_xref="GI:20152849"

/translation="WDSPJPSLSFLVILFIALIKSCQCPHEKRLNHLSTYNTL
ERVAENSEPLEVGLQTLQIDVDENKQQLINMLSEWDNYDNLRWNETSYGVK
DLRIATNKPKVLPMYNSADGSGQVHTNIVVKHSGSLVPPGFKSCMDITD
FPEDQCEKMGFSYQNGNLDVNSDGDGSLDTNGEWLLAMPKOKTIVIA
CSEPVEDVFTTQIRRRYQLLVNLPVCLVLSMGLLFTPDGSKGLTVLIT
LSUTLVNVAETLPQVSDAPELLGTFCNFMVAGSVLTVKVNHLRTADIDH
PMUKSVFLQWLPWLMRGPGRKTRIKTLLSNMRKEALLKRSKSLNLDVDD
DFRHTJSGSTALIGSSASGPTVEEHTTATQCNHKLHLLKELQFIYARWKADD
EAEILIGFEAAVMVDFCLIVTEFTTATVTVLLSAPHILVO"

variation

```

134 /gene="nAcRalpha-30D"
135 /note="results in glycine to serine substitution; compared
136 to the sequence deposited in GenBank Accession Number
137 AE003626"
138 /accession="n3"

```

variation

```

846 /gene="nAcRalpha-30D"
      /note="results in asparagine to serine substitution;
      compared to the sequence deposited in Genbank Accession
      Number AE003626"
      /replace="na"

```

ORIGIN

Alignment Scores:		
Pred. ID:	5.11e-163	Length:
Score:	1792.50	Matches:
Percent Similarity:	78.88%	Conservative:
Best Local Similarity:	67.93%	Mismatches:
Query Match:	67.26%	Indels:
DB:	3	Gaps:
		5
US-09-303-232-4 (1-496) x AE321447 (1-2023)		


```
QY 287 LeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValProLeuLeuGlyThr 306
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
890 CTAACCTTGTCCCGAGTCCATGCCACACGTCGGATGCTGCTCTTATAGGCACC 949
QY 307 TyrPheAsnCysIleMetPheMetValAlaSerSerValValSerThrIleLeuIleLeu 326
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
950 TACTTCAATTGCATCATGTTTCATGTCCTCGTGGTGGTGCTGACAGTAGTGGTCTC 1009
QY 327 AsnTyrHisHisArgHisAlaAspThrHisGluMetSerAspTrpIleArgCysValPhe 346
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1010 AACTACCACCATCGCACAGCGGACATTCACGAGATGCCACCGTGGATCAAGTCGGTTTC 1069
QY 347 LeuTyrTrpLeuProTtpValLeuArgMetSerArgProGlySerAlaThrThrProPro 366
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1070 CTACATGGTGGCTGGATGCTTGGCAATGGTGGACCGCGTGGCAGATTACACGCAA 1129
QY 367 ProAlaArgValProProProProAspLeuGluLeuArgGluArgSerSerLysSer 385
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1130 ACAATACTATTAAAGCAATCGCATGAGAGCTGGAGCTAAGAGCGCTCTCCAAATCC 1189
QY 386 LeuLeuAlaAsnValLeuAspIleAspAspPheArgHisProGlnAlaGlnGlnPro 405
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1190 CTGCTGGCCAAATGCTCTCGACATCGACGACACTTCGGGCACACAAATATCT- 1240
QY 406 GlnCysCysArgTyrTyrArgGlyGlyGluGluAsnGlyAlaGlyLeuAlaAla- 423
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1241 -----GGTCCCAACCCCATTTGGCTCGTGGCCAGCTTC 1276
QY 424 -----HisSerCysPheGlyValAspTyr---GluLeu 433
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1277 GGTCCGCCCAACACGGTGGAGGAGCATCACACGCCCATCGCTCCATCACAAGATCTT 1336
QY 434 SerLeuIleLeuLysGluIleArgValIleThrAspGlnMetArgLysAspAspGluAsp 453
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1337 CATCTAATTCTCAAGAAATTGCAATTTATTACGGCGCGGATGGCAAAAGCTGACACGAA 1396
QY 454 AlaAspIleSerArgAspTrpLysPheAlaAlaMetValValAspArgLeuCysLeuIle 473
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1397 GCGAATGTATCGCGGATTTGGAAGTTTCGGGCAATGGTTGTGGATAGATTTGTTTAATT 1456
QY 474 IlePheThrLeuPheThrIleIleAlaThrLeuAlaValLeuLeuSerAlaProHisIle 493
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1457 GTTTTCACGCTCTTCAGGATTATTGCAACGGTTACGGTGTGCTCTCCGCTCCGCACATA 1516
QY 494 MetVal 495
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1517 ATCGTG 1522
```

Search completed: May 9, 2004, 00:55:05

Job time : 5495.09 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2004, 06:39:04 ; Search time 520.982 Seconds
(without alignments)
4044.486 Million cell updates/sec

Title: US-09-303-232-4
Perfect score: 2665
Sequence: 1 MCGRRARRSHLAAPAGLLLL.....LFTIATLAVLLSAPHIMVS 496

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09303232/runat_07052004_101109_23873/app_query.fasta_1.2261
-DB=N_Geneseq_29Jan04 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09303232_@CGN_1_1200_@runat_07052004_101109_23873 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002s:*
7: Geneseq2003as:*
8: Geneseq2003bs:*
9: Geneseq2003cs:*
10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2665	100.0	3700	3 AAZ24476	Aaz24476 H. viresc
2	1850	69.4	2886	3 AAZ24475	Aaz24475 D. melano
3	1809	67.9	3109	3 AAZ24477	Aaz24477 H. viresc
4	1623.5	60.9	1540	4 ABL07231	AbL07231 Drosophil
5	1417.5	53.2	936	4 ABL13733	AbL13733 Drosophil
6	1267	47.5	2769	2 AAT59196	Aat59196 Human neu
7	1247	46.8	1876	2 AAV12197	Aav12197 Human neu
8	1247	46.8	1876	2 AAT48239	Aat48239 Neuronal

9	1247	46.8	1876	6 ABS54875	AbS54875 Human neu
10	1247	46.8	1876	6 ABV73248	Abv73248 Human neu
11	1247	46.8	1876	8 ADA10864	Ada10864 Human neu
12	1246	46.8	1509	3 AAC58195	Aac58195 Human PRO
13	1246	46.8	1509	4 AAC90380	Aac90380 Wild-type
14	1242	46.6	1590	2 AAV44687	Aav44687 V274T var
15	1240	46.5	1509	4 AAC90385	Aac90385 Mutant hu
16	1236	46.4	1509	4 AAC90386	Aac90386 Mutant hu
17	1230	46.2	1509	4 AAC90387	Aac90387 Mutant hu
18	1219.5	45.8	2101	2 AAT59197	Aat59197 Neuronal
19	1215.5	45.6	1964	6 ABZ11298	Abz11298 Human pol
20	1193	44.8	803	4 ABL07799	AbL07799 Drosophil
21	1057.5	39.7	1915	4 AAD20962	Aad20962 Caenorhab
22	1011	37.9	5096	4 ABL13732	AbL13732 Drosophil
23	1003.5	37.7	1416	4 AAC90382	Aac90382 Chimeric
24	968	36.3	1896	6 AAL45867	Aal45867 Modified
25	960	36.0	1908	2 AAT48236	Aat48236 Neuronal
26	960	36.0	1908	8 ADA10856	Ada10856 Human neu
27	960	36.0	2277	2 AAV12199	Aav12199 Human neu
28	960	36.0	2277	2 AAQ90387	Aaq90387 Alpha 2 s
29	960	36.0	2277	6 ABS54870	AbS54870 Human neu
30	960	36.0	2277	6 ABV73243	Abv73243 Human neu
31	960	36.0	2664	2 AAT48235	Aat48235 Neuronal
32	960	36.0	2664	6 ABK92165	Abk92165 Prostate
33	960	36.0	2664	8 ADA10854	Ada10854 Human neu
34	960	36.0	2664	9 ADC71170	Adc71170 Human 205
35	960	36.0	2666	9 ADC71168	Adc71168 Human 205
36	957.5	35.9	1869	6 AAL45873	Aal45873 Modified
37	957.5	35.9	1869	6 ABL54794	AbL54794 Modified
38	954	35.8	1584	7 ADA83809	Ada83809 Human CHR
39	935.5	35.1	2210	4 ABL10275	AbL10275 Drosophil
40	929.5	34.9	1743	2 AAT48232	Aat48232 Neuronal
41	929.5	34.9	1743	6 ABS54874	AbS54874 Human neu
42	929.5	34.9	1743	6 ABV73247	Abv73247 Human neu
43	929.5	34.9	1743	8 ADA10862	Ada10862 Human neu
44	923.5	34.7	2385	4 ABL11821	AbL11821 Drosophil
45	913	34.3	1915	2 AAT48241	Aat48241 Neuronal

ALIGNMENTS

RESULT 1

ID AAZ24476 standard; CDNA to mRNA; 3700 BP.

XX AAZ24476;

DT 17-FEB-2000 (first entry)

DE H. virescens acetyl-choline receptor DNA from clone Hva7-1.

KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
KW neurotransmission; plant protection agent; conductance; ACHR; ds.

OS Heliothis virescens.

FH Key Location/Qualifiers

CDS 335..1825

FT /*tag= a

FT /product= "acetyl-choline receptor"

PN DB19819829-Al.

XX 11-NOV-1999.

XX 04-MAY-1998; 98DE-01019829.

XX 04-MAY-1998; 98DE-01019829.

XX (FARB) BAYER AG.

XX Adamczewski M, Oellers N, Schulte T;

DR WPI: 2000-014207/02.
 DR P-PSDB; AAY50815.
 XX New nucleic acid encoding a nicotinic acetylcholine receptor from
 PT insects, used to identify potential insecticides.
 XX
 XX Claim 1a; Page 14-17; 26pp; German.
 PS
 XX This invention describes a novel nucleic acid (NA) encoding a nicotinic
 CC acetylcholine receptor (I) from insects which can be used as an
 CC insecticide. Inhibitors of (I) interfere with neurotransmission. (I)
 CC (also vectors containing it, its regulatory regions, and antibodies
 CC directed against (I)-encoded proteins) are used to screen for: (a) plant
 CC protection agents that alter conductance of ACHR, potentially useful as
 CC insecticides, or (b) genes which encode polypeptides that are involved in
 CC formation of functionally related ACHR in insects. (I) are also used to
 CC isolate and characterize the specified regulatory regions and for
 CC recombinant production of (II). This sequence encodes an acetylcholine
 CC receptor isolated from *Heliothis virescens*
 XX
 XX Sequence 3700 BP; 893 A; 953 C; 944 G; 910 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.72e-257	Length:	3700
Score:	2665.00	Matches:	496
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-09-303-232-4 (1-496) x AA224476 (1-3700)

Qy	1	MetGlyGlyArgAlaArgSerHisLeuAlaAlaProAlaGlyLeuLeuLeuLeu	20
Db	335	ATGGCGGGCGGGCGCGCGCTCGACATTGGCGGGCGCGCGCGCTCGTGTGCTG	394
Qy	21	CysLeuLeuTrpProArgGlyAlaArgCysGlyTyrHisGluLysArgLeuHis	40
Db	395	TGCTGTCTTGCGCGGGGGGCGCGCTCGCGGTACCGAGAGCGGCTACTGCGAC	454
Qy	41	LeuLeuAspHisTyrAsnValLeuGluArgProValValAsnGluSerAspProLeu	60
Db	455	CTATTGGACCACTCAACAGTACTGGAGAGCGCGCTCGTCAACAGAGCGCGCG	514
Qy	61	LeuSerPheGlyLeuThreuleuMetGlnIleleAspValAspGluLysAsnGln	80
Db	515	CTCTCTTTCGGCTCACGCTCATGAGATCATCGACGTGAGAGAGAGAGAGAG	574
Qy	81	IleThrAsnIleTyrLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsn	100
Db	575	ATAACAACATCTGGCTAACTAGATGAGTGAATGATGAATTTGAGGTGGAC	634
Qy	101	AspPheGlyValLysAspLeuArgValProProHisArgLeuTrpLysProAsp	120
Db	635	GATTTTCGGCGGGCTCAAGATTTAAGAGTGCACCCACAGACTATGGAAC	694
Qy	121	LeuMetTyrAsnSerAlaAspGluGlyPheAspSerThrTyrProThrAsnVal	140
Db	695	CTTATGTACAACAGCGCGGACGAAGGGTTTCGACAGACGATATCCCAAC	754
Qy	141	ArgAsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSerThrCys	160
Db	755	CGGAACAACGGCTCGTGTCTGTACGTGCGCGCGCGCATCTTCAAGAGCA	814
Qy	161	AspIleThrTrpPheProPheAspAspGlnArgCysGluMetLysPheGlySer	180
Db	815	GACATCACCTGGTTCCCTTCGACGACCAACGATGCGAGATGAAGTTTGG	874
Qy	181	TyrAspGlyTyrGlnLeuAspLeuGlnLeuAspGluGlyGlyAspIleSerSer	200
Db	875	TATGATGGTTATCAGTTGGATCTACAACTACAGAGATGAAGGGGGGAGAT	934
Qy	201	PheValThrAsnGlyGluTrpGluLeuIleGlyValProGlyLysArgAsnGlu	220

Db	935	TTTGTACGAATGGGAATGGAGTTAATAGAGTCCCGCGCAAGCGCAACGAGAT	994
Qy	221	TyrAsnCysCysProGluProTyrIleAspIleThrPheAlaValValIleArg	240
Db	995	TACAACTGTTTTCGGAGGCCATACATCGATCAGTTTGGTGGTGGTATCGG	1054
Qy	241	ThrLeuTyrTyrPhePheAsnLeuValProCysValLeuIleAlaSerMet	260
Db	1055	ACGCTCTACTACTTCTTCAATCTGATGTCCTCGGTGCTCATCGCCTCAT	1114
Qy	261	LeuGlyPheThrLeuProAspSerGlyGlyLysLeuSerLeuGlyValThr	280
Db	1115	TTGGGTTTCACTTTCCTCCAGACTCCGAGAGAAAAGTTGTCTTAGGTGT	1174
Qy	281	LeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAlaThr	300
Db	1175	CTGCTGTTGACGGTTCCTCAACATGGTGGCGAGACGATGCCAGCAGCG	1234
Qy	301	ValProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAlaSer	320
Db	1235	GTGCTTGTCTGGCACCCTACTTCACTGCTCATGTTTCTTCTTCTTCT	1294
Qy	321	SerThrIleLeuIleLeuAsnTyrHisHisArgHisAlaAspThrHisGlu	340
Db	1295	TCCACCATCTGATCTCACTACCAACCGCGCAGCAGACACTCAGAAATG	1354
Qy	341	TrpIleArgCysValPheLeuTyrTrpLeuProTrpValLeuArgMetSer	360
Db	1355	TGGATTCTGCTGCTTCTTCTTATTTGCTGCGTGGTGGTGGTGGTGGT	1414
Qy	361	SerAlaThrTrpProProAlaArgValProProProAspLeuGluLeuArg	380
Db	1415	TGGCGGACGACGCGCGCGCGCGCTACTCTCGCGCGCGGACCTGGAG	1474
Qy	381	ArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArg	400
Db	1475	CGTCTCTCAAGTGTCTCTAGCGAAGCTGTGACATCGATGACGACTT	1534
Qy	401	GlnAlaGlnGlnProGlnCysArgTyrTyrArgGlyGlyGluAsnGlyAla	420
Db	1535	CAAGCGCAGCAGCGCAATGCTCGCATACTACAGGGGGGGGAGAGAA	1594
Qy	421	LeuAlaAlaHisSerCysPheGlyValAspTyrGluLeuSerLeuIleLys	440
Db	1595	TTGGCGGCGCAGCTTGTCTCGTGTGCTACTACGAGCTCTCCTCAT	1654
Qy	441	ArgValIleThrAspGlnMetArgLysAspAspGluAspAlaAspIleSer	460
Db	1655	AGAGTCATCAGATCAGATGCGCAAGGACGACGAGATGCGGACATTT	1714
Qy	461	LysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrLeu	480
Db	1715	AAGTTCGGCGCGCATGGTGTGGACAGCTGTGCTTATTTATTTACCT	1774
Qy	481	IleAlaThrLeuAlaValLeuSerAlaProHisIleMetValSer	496
Db	1775	ATCGCACGCTAGCGTGTGCTGTCGCGCGCACACATCATGTGTG	1822

RESULT 2

AAZ24475
 ID AAZ24475 standard; cDNA to mRNA; 2886 BP.

XX AAZ24475;
 AC AAZ24475;

XX 17-FEB-2000 (first entry)

DE D. melanogaster acetylcholine receptor DNA from clone Da7.

KW Acetylcholine receptor; nicotinic; insect; insecticide; screening;
 KW neurotransmission; plant protection agent; conductance; AChR; ds.

XX Drosophila melanogaster.


```
Db 890 ATTCTTAACCTTCTCGCGAGTCCATGCGCACAGCTCGGATGCTGTCTCTTATAGG 949
Qy 305 yThr----- 306
Db 950 TAC-ACACACAGCATATTAACACGCTATAGCTTAAAGTACGACCGATCGGTTCGACTAT 1008
Qy 307 -----TyrPheAsnCysIleMetPheMetVa 315
Db 1009 CGACCACTACTACAGTGTGAGCAGAAAGCACCTACTTCAATTGCATCATGTTCAUGGT 1068
Qy 315 lalaSerSerValValSerThrIleLeuIleLeuAsnTyrHisHisArgHisAlaAspTh 335
Db 1069 CGCCTCGTGGTGTGTCAGAGTAGTGTCTCACTACCATCGCACAGCGGACAT 1128
Qy 335 rHisGluMetSerAspTTPlleArgCysValPheLeuTyrTrpLeuProTrpValLeuAr 355
Db 1129 TCACGAGATGCCACGCTGGATCAAGTCGCTTTTCTCAATGGCTGGCTGGATCTGCG 1188
Qy 355 gMetSerArgProGlySerAlaThrThrProProAlaArgValPro---ProProPr 374
Db 1189 AATGGTCCAGCCGGTCCAGATTACAGCAAAACATACTATTAAAGCAATCGCATGAA 1248
Qy 374 oAspLeuGluLeuArgGluArgSerSerIysSerLeuLeuAlaAsnValLeuAspIleAs 394
Db 1249 GGAGCTGGAGCTAAAGGAGCGCTCTCCAAATCCCTGCTGGCCAATGTCTCGACATCGA 1308
Qy 394 pAspAspPheArgHisProGlnAlaGlnGlnProGlnCysCysArgTyrTyrArgGlyGl 414
Db 1309 CGACCACTTCGGGCACAAATATCT-----GG 1335
Qy 414 yGluGluAsnGlyAlaGlyLeuAlaAla----- 423
Db 1336 CTCCAAACCGCATTTGGCTGCTCGCCAGCTTCGGTCGCCCAACACGGTGGAGGACGA 1395
Qy 424 -HisSerCysPheGlyValAspTyr---GluLeuSerLeuIleLeuLysGluIleArgVa 442
Db 1396 TCACACGGCCATCGCTGCAATCACAAAGATCTTCAATCTCAAAGAATTGCAATT 1455
Qy 442 lIleThrAspGlnMetArgLysAspAspGluAspAlaAspIleSerArgAspTrpLysPh 462
Db 1456 TATTACGGCGCGGATGCGCAAGAGCTGACGACGAAGCAATTGATCGCGGATTGGAAGTT 1515
Qy 462 eAlaAlaMetValValAspArg 469
Db 1516 CGCGCAATGTTTGGATAGG 1537
RESULT 5
ID ABL13733
AC ABL13733;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 35681.
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX (PEKE ) PE CORP NY.
XX PA
XX
```

```
PI Venter JC, Adams M, Li PWD, Myers EW;
DR WPI: 2001-656860/75.
DR P-FSDB; ABB69630.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX Claim 1; SEQ ID NO 35681; 2lpp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins (ABBS7737-
CC ABBS7072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 936 BP; 238 A; 213 C; 239 G; 246 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.23e-132 Length: 936
Score: 1417.50 Matches: 265
Percent Similarity: 91.78% Conservative: 14
Best Local Similarity: 87.17% Mismatches: 16
Query Match: 53.19% Indels: 9
DB: 4 Gaps: 2
US-09-303-232-4 (1-496) x ABL13733 (1-936)
Qy 2 GlyGlyArgAlaArgArgSerHisLeuAlaProAlaGlyLeuLeuLeuLeuLeuCys 21
Db 46 GGTGGCAGAATG-----CTGTCATATGCGCTGGGACTTTTATTATGATA --- 90
Qy 22 LeuLeuTrpProArgGlyAlaArgCysGlyTyrHisGluLysArgLeuLeuHisLeu 41
Db 91 -----CCGGCTTGTGGCGCTGACCCCATGAGAGCGCTACTCCACGCCCTT 138
Qy 42 LeuAspHisTyrAsnValLeuGluArgProValValAsnGluSerAspProLeuLeu 61
Db 139 CTGGACAACTACACAGCCTGGAGCGTCGGTGGTCAATGAATCCGATCCATTCGCAACTG 198
Qy 62 SerPheGlyLeuThrLeuMetGlnIleIleAspValAspGluLysAsnGlnLeuLeu 81
Db 199 AGCTTCGGACTAACACTCATGCAGATTATCGATGTGGACGAAAGAAATCACTGCTTATA 258
Qy 82 ThrAsnIleTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAsp 101
Db 259 ACGAATATTGGCTCAAAATGGAAATGGAACGATATGATATCTTCGATGGAATTCGAGTGAG 318
Qy 102 PheGlyGlyValLysAspLeuArgValProProHisArgLeuTrpLysProAspValLeu 121
Db 319 TTCGTGTGTGTGGGATCTCGCAATTCGCGACATCGCTATGGAACCGGATGACTAGT 378
Qy 122 MetTyrAsnSerAlaAspGluGlyPheAspSerThrTyrProThrAsnValValValArg 141
Db 379 ATGTACACAGTGGCGACGAGGGCTTCGATGGAACGTCAGCCACAAATGTGGTTCGC 438
Qy 142 AsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSerThrCysLysIleAsp 161
Db 439 AATAATGGAGCTGTCTGTACTACCGCAGGTATATTATTAAGTCAACGTTGAAGATCGAC 498
Qy 162 IleThrTrpPheProPheAspAspGlnArgCysGluMetLysPheGlySerTrpThrTyr 181
Db 499 ATTACGTGTGTTCATTTCGACATCAGATGTGAAATGAAATTTGGTTCTGGACCTAC 558
Qy 182 AspGlyTyrGlnLeuAspLeuGlnLeuGlnAspGluGlyGlyAspIleSerSerPhe 201
Db 559 GATGGGTTTCAGTTGGACCTGTCAGTTGCAGGACGAAGCTGGTGGCGACATTCTTAGCTTT 618
```


QY 261 LeuGlyPheThrLeuProProAspSerGlyGluLeuSerLeuGlyValThrIleLeu 280
DB 818 TTAGCTCTTCGCTCAGCAGATCTCAGGAGAAAGATCTCTAGGTATACAGTTTA 877
QY 281 LeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAla 300
DB 878 TTGTCTCTCAGCGTTCATGTACTCTGCTGCTGAAATTAATGCGAGCAACATCTGATTCT 937
QY 301 ValProLeuLeuGlyThrTyPheAsnCysIleMetPheMetValAlaSerSerValVal 320
DB 938 GTGCCCTTAATGTCTAGTATTTTCCACACCATGATTAATTTGGCTCTCTGTGTTT 997
QY 321 SerThrIleLeuLeuAsnTyHisHisArgHisAlaAspThrHisGluMetSerAsp 340
DB 998 GTCACTGTTATCGTTTCAATACCATCATCAGATCCAGATGGGGGAAATGCGTAAA 1057
QY 341 TrpIleArgCysValPheLeuTyTrpLeuProTrpValLeuArgMetSerArgProGly 360
DB 1058 TGGACAGAGTCATCTCTGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1117
QY 361 SerAlaThrProProAlaArgValProProProProProProProProProProPro 380
DB 1118 GAA-----GATAAGTGGTCCCGCTGTCAACATAAAGAGCGCGA 1159
QY 381 ArgSerSerLysSerLeuLeuAlaAsnValLeuAsp----- 392
DB 1160 TGCAGCTGTCAAGCATGAGATGACACTGTGAGTGTGTCAGCAATGAGTAATGGGAAC 1219
QY 393 -----IleAspAspPheArgHisProGlnAlaGlnGln 404
DB 1220 ATGCTGTATATTGGTTCGAGGGCTGATGGGTTCACCTGCACACCCACCACTGATTCA 1279
QY 405 ProGlnCysCysArgTyTrpArgGly-----GluGluAsnGlyAlaGlyLeuAla 422
DB 1280 GGGGTGATCTGGGAGGATGACCTGTTCAACAGAGGAGAAATCTTCTGCACAGT 1339
QY 423 AlaHisSerCysPheGlyValAspTyTrpGluLeuSerLeuLeuLeuLeuLeuLeuVal 442
DB 1340 GGCACCCCTCTGAGGCG---GACCAGATTGGTGAATCTTGAAGAGTTCAGATAC 1396
QY 443 IleThrAspGlnMetArgLysAspAspGluAspAlaPheSerArgAspTrpLysPhe 462
DB 1397 ATTGCAACAGGTTTCAGAGACCCAGGATGAAGAAGACCAATTTGCAACCAATGGAAGTT 1456
QY 463 AlaAlaMetValValAspArgLeuCysLeuIleIlePheThrLeuPheThrIleIleAla 482
DB 1457 GCAGCTCTGTAGTAGTCGGCTCTGTGTGATGGCATTTTCGGTCTTCCACATCTTGT 1516
QY 483 ThrLeuAlaValLeuSerAlaProHisIleMet 494
DB 1517 ACAATTGGCATCTTAATGTCAGCACCAAACTTTGTA 1552

RESULT 7

ID AAV12197 standard; cDNA; 1876 BP.
XX AC AAV12197;
XX DT 14-MAY-1998 (first entry)
DE Human neuronal nicotinic acetylcholine receptor alpha-7 subunit cDNA.
XX Human neuronal nicotinic acetylcholine receptor; alpha-7 subunit;
KW brain tissue; screening; NACHR; antibody; ds.
XX OS Homo sapiens.
XX FH Key
XX CDS 73..1581
XX FT Location/Qualifiers
XX FT *tag= a
XX FT /product= "neuronal nicotinic acetylcholine receptor
XX FT alpha-7 subunit"

PN W09420617-A2.
XX 15-SEP-1994.
XX 08-MAR-1994; 94WO-US002447.
XX 08-MAR-1993; 93US-00028031.
XX (SALK) SALK INST BIOTECHNOLOGY IND ASSOC.
XX Elliott KJ, Ellis SB, Harpold MM;
XX WPI; 1994-303024/37.
XX P-PSDB; AAW44153.
XX Human neuronal nicotinic acetylcholine receptor subunits and DNA - also
XX transformed cells useful for screening cpds. which modulate activity of
XX the receptor.
XX Claim 8; Page 78-79; 99pp; English.
XX The present sequence encodes a human neuronal nicotinic acetylcholine
XX receptor (NACHR) subunit. The cells expressing the alpha and/or beta
XX NACHR subunits may be used in a method of screening compounds to identify
XX any which modulate the activity of human neuronal NACHR. Subunit specific
XX antibodies may be used to monitor the distribution and expression density
XX of various subunits in normal vs diseased brain tissues. Testing of
XX single receptor subunits or specific receptor subunit combinations with a
XX variety of potential agonists or antagonists provides information with
XX respect to the function and activity of the individual subunits and
XX should lead to the identification and design of compounds that are
XX capable of very specific interaction with one or more receptor subtypes.
XX The resulting drugs should exhibit fewer unwanted side effects than drugs
XX identified e.g. screening with cells that express a variety of subtypes
XX SQ Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 4.91e-115 Length: 1876
Score: 1247.00 Matches: 247
Percent Similarity: 60.47% Conservative: 85
Best Local Similarity: 44.99% Mismatches: 129
Query Match: 46.79% Indels: 88
DB: 2 Gaps: 8

US-09-303-232-4 (1-496) x AAV12197 (1-1876)

QY 3 GlyArgAlaArgArg-----SerHisLeuAlaAla--- 12
DB 1 GGGCCGAGCGCGAGCGCGCGGCGAGACGTGGAGCGCGCGCTCGCTCGCAGCT 60
QY 13 -----ProAlaGlyLeuLeuLeuLeuLeuLeuLeuTrp 24
DB 61 CGGGAGCTCAACATCGCTGCTCGCGGAGCGCTGCTGGCTGGCGCTCGCTC 120
QY 25 ProArgGlyAlaArgCysGlyTyHisGluLysArgLeuLeuHisHisLeuLeuAspHis 44
DB 121 CTGCACGTCTCCCTGCAAGCGGAGTTCAGAGAGAGCTTTACAGAGCTGGTCAAGAAC 180
QY 45 TyrAsnValLeuGluArgProValValAsnGluSerAspProLeuGlnLeuSerPheGly 64
DB 181 TACAATCCCTTGGAGAGCGCGCTGGCCAAATGACTGCACCACTCACCTCTACTTCTCC 240
QY 65 LeuThrLeuMetGlnIleIleAspValAspGluLysAsnGlnLeuLeuLeuLeuLeuLeu 84
DB 241 CTGAGCTCTCTGAGTATGAGTGGATGAGAGAACCAAGTTTAAACCACCAACATT 300
QY 85 TrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspPheGlyGly 104
DB 301 TGGCTGCAATGTCTTGGACAGATCACTATTATACAGTGGAAATGTGTCAAGATATCCAGG 360
QY 105 ValLysAspLeuArgValProHisArgLeuTrpLysProAspValLeuMetTyAsn 124

ID XX ABV73248 standard; cDNA; 1876 BP.
AC XX ABV73248;
DT XX 22-JAN-2003 (first entry)
DE XX Human neuronal NACHR alpha7 subunit encoding cDNA.
XX KW Human; neuronal; nicotinic acetylcholine receptor; NACHR; drug screening;
KW KW immunochemistry; NACHR alpha7 subunit; gene; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 73..1581
XX FT /*tag= a
XX FT /product= "NACHR alpha7 subunit"
XX FT /note= "neuronal nicotinic acetylcholine receptor"
XX XX
XX XX WO200259266-A2.
XX XX 01-AUG-2002.
XX XX 29-OCT-2001; 2001WO-US050985.
XX XX 01-NOV-2000; 2000US-00703951.
XX XX (MERI) MERCK & CO INC.
XX XX Gillespie A, Claeps BO, Chavez-Noriega LE, Siegel R, Elliott KJ;
XX XX WPI; 2002-698532/75.
XX XX P-PSDB; ABB82435.
XX XX Cell comprising nucleic acids encoding human alpha and beta subunits of
XX XX neuronal nicotinic acetylcholine receptors, useful for in vitro screening
XX XX of a drug substance in a test system specific for humans.
XX XX Example; Page 129-130; 143pp; English.
XX XX The invention relates to a suitable host cell transfected with an
XX XX isolated nucleic acid molecule comprising a sequence of nucleotides or
XX XX ribonucleotides that encodes at least one alpha or beta subunit of a
XX XX human neuronal nicotinic acetylcholine receptor (NACHR). The compositions
XX XX and methods of the present invention, which provide a means to prepare
XX XX synthetic or recombinant receptors and receptor subunits that are
XX XX substantially free of contamination from many other receptor proteins,
XX XX are useful for observing the effect of a drug substance on a particular
XX XX subtype to perform in vitro screening of the drug substance in a test
XX XX system that is specific for humans. The antibodies can be used in
XX XX immunochemistry and for diagnostic and therapeutic applications. The
XX XX present sequence represents a human neuronal NACHR alpha7 subunit
XX XX encoding cDNA
XX XX Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;
XX XX
XX XX Alignment Scores:
XX XX Pred. No.: 4,91e-115 Length: 1876
XX XX Score: 1247.00 Matches: 247
XX XX Percent Similarity: 60.47% Conservative: 85
XX XX Best Local Similarity: 44.99% Mismatches: 129
XX XX Query Match: 46.79% Indels: 88
XX XX DB: 6 Gaps: 8
XX XX
XX XX US-09-303-232-4 (1-496) x ABV73248 (1-1876)
XX XX
XX XX 3 GlyArgAlaArg-----SerHisLeuAlaAla--- 12
XX XX |||||
XX XX 1 GCGCGAGCGCGAGCGCGCGGCGACAGCGAGAGCGTGGAGCGCGCGGCTCGTGCAGCT 60
XX XX |||||
XX XX 13 -----ProAlaGlyLeuLeuLeuLeuLeuLeuTrp 24
XX XX |||||
XX XX 61 CCGGGACTCAACATGCGTCTCGCGGAGGAGCGTCTGGCTGCGCGCGTGGCTC 120

QY 25 ProArgGlyAlaArgCysGlyTyrHisGluLysArgLeuLeuHisHisLeuLeuAspHis 44
Db 121 CTGCACGTGTCCTCGCAGGCGAGTTCACAGGAAGCTTTACAGGAGCTGGTCAAGAAC 180
QY 45 TyrAsnValLeuGluArgProValValAsnGluSerAspProLeuGlnLeuSerPheGly 64
Db 181 TACAATCCCTTGGAGAGCGCGTGGCCAATAGATCGCAACCACTCACCGCTACTCTTCC 240
QY 65 LeuThrLeuMetGlnIleAspValAspGluLysAsnGlnLeuLeuIleThrAsnIle 84
Db 241 CTGAGCCTCTGTCAGATCATGGATGAGAGAACCAAGTATTAAACCAACCAAT 300
QY 85 TrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspPheGly 104
Db 301 TGGCTGCAATGTCTTGGACAGATCACTATTATACAGTGAATGTGTGAGATATCCAGG 360
QY 105 ValLysAspLeuArgValProHisArgLeuTrpLysProAspValLeuMetTyrAsn 124
Db 361 GTGAAGACTGTTCGTTTCCAGATGGCAGATTGGAAACCAAGACATTTCTCTATAAC 420
QY 125 SerAlaAspGluGlyPheAspSerThrTyrProThrAsnValValValArgAsnAsnGly 144
Db 421 AGTGCTGATGAGCGCTTTGACCGCACATTCACACTACGTGTGGTGAATCTCTCTGG 480
QY 145 SerCysLeuTyrValProGlyIlePheLysSerThrCysLysIleAspIleThrTrp 164
Db 481 CATTCGCCAGTACCTCCAGGCATATTCAAGAGTTCCTGCTACATCATGATGACGCTGG 540
QY 165 PheProPheAspAspGlnArgCysGluMetLysPheGlySerTrpThrTyrAspGlyTyr 184
Db 541 TTTCCCTTTGATGTGCAGCACTGCAAACTGAAGTTGGTCTCTGCTGTACGAGGCTGG 600
QY 185 GlnLeuAspLeuGlnLeuAspGluGlyGlyAspIleSerSerPheValThrAsn 204
Db 601 TCCTTGGATCTGCAGATGACGAG-----GCAGATATACGTGGGTATATCCCCAAT 651
QY 205 GlyGluTrpGluLeuIleGlyValProGlyLysArgAsnGluIleTyrTyrAsnCysCys 224
Db 652 GGAGATGGGACCTAGTGGGAATCCCGGCAAGAGGAGTGAAGGTTCTATGAGTGTCTG 711
QY 225 ProGluProTyrIleAspIleThrPheAlaValValIleArgArgLysThrIleTyr 244
Db 712 AAAGAGCCCTACCCCGATGTCACTTCACGTGACCATGCGCGCGAGGACGCTCTACAT 771
QY 245 PhePheAsnLeuIleValProCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThr 264
Db 772 GGCCTCAACCTGCTGATCCCTGTGTGTCTATCTCCGCCCTCGCCTGCTGTGTCTCTG 831
QY 265 LeuProProAspSerGlyLysLeuSerLeuGlyValThrIleLeuLeuSerIleThr 284
Db 832 CTTCTCGCAGATTCCGGGGAGAAAGATTCCCTGGGGATAACAGATCTTACTCTCTTACC 891
QY 285 ValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValProLeuLeu 304
Db 892 GTCTTCATGCTGCTGCTGCTGAGATCATGCCCGCAACATCCGATTCGATGACCATGATA 951
QY 305 GlyThrTyrPheAsnCysIleMetPheMetValAlaSerSerValValSerThrIleLeu 324
Db 952 GCCAGTACTTCGCCAGCACCATGATCATCGTGGGCCCTCTCGTGTGTGACGCTGATC 1011
QY 325 IleLeuAsnTyrHisHisArgHisAlaAspThrHisGluMetSerAspTrpIleArgCys 344
Db 1012 GTGCTGCAGTACCAACCAACCAACCGCGGGGCAAGATCCCAAGTGGACCAAGATC 1071
QY 345 ValPheLeuTyrTrpLeuProTrpValLeuArgMetSerArgProGly----- 360
Db 1072 ATCTTCTGAATGTGTGCGCGTGGTCTCTSCGATGAAGAGCGCGCGGAGGACAGAGTG 1131
QY 361 -----SerAla 362
Db 1132 CGCCCGGCTGCCAGCAAGCAGCGGCGCTGCAGCGCTGGCAGTGTGGAGATGACGCGC 1191

Percent Similarity:	62.45%	Conservative:	86
Best Local Similarity:	45.98%	Mismatches:	12
Query Match:	46.75%	Indels:	70
DB:	3	Gaps:	6

US-09-303-232-4 (1-496) x AAC58395 (1-1509)

QY 12 AlaProAlaGlyLeuLeuLeuLeuCysLeuLeuTrpProArgGlyAlaAryCysGly 31
Db : : : : | | | | : : : :
10 TCGCCGGAGGCGCTTGGCGGTGGCGGCTGCTCTGCACGTGCCCTCAAGGC 69

QY 32 TyrHisGluLysArgLeuLeuHisLeuLeuAspHisTyrAsnValLeuGluArgPro 51
Db : : : : | | | | : : : :
70 GAGTTCCAGAGAGACTTTACAAGGAGCTGGTCAAGAATACTAAATCCCCTTGAGAGGCC 129

QY 52 ValValAsnGlnSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeu 71
Db | | | | : : : :
130 GTGGCCAATGACTCGCAACCACTCACCGCTACTTCTCCCTGAGCCTCTCGCATCATG 189

QY 72 AspValAspGluLysAsnGlnLeuLeuLeuThrAsnIleTrpLeuLysLeuGluTrpAsn 91
Db | | | | : : : :
190 GACGTGGATGAGGAAGAACCAAGTTTAACCAACAACATTGGCTGCAATGTCTGGACA 249

QY 92 AspMetAsnLeuArgTrpAsnThrSerAspPheGlyGlyValLysAspLeuArgValPro 111
Db | | | | : : : :
250 GATCACTATTACAGTGAATGTGTGAGAAATATCCAGGGGTGAAGACTGTCCTGTTCCCA 309

QY 112 ProHisArgLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp 131
Db : : : : | | | | : : : :
310 GATGCCAGATTGGAAACCAGACATTCTTCTATAACAGTGTGATGAGCGCTTTGAC 369

QY 132 SerThrTyrProThrAsnValValArgAsnAsnGlySerCysLeuTyrValProPro 151
Db : : : : | | | | : : : :
370 GCCACATCCACACTAACGTGTGTGTGAATCTTCTGGGCATTGGCAGTACTCGCTGCTCA 429

QY 152 GlyIlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArg 171
Db | | | | : : : :
430 GGCATATTCAAGAGTTCCTGTACATCATGATGATCGCTGGTTTCCCTTTGATGTGCAGCAC 489

QY 172 CysGluMetLysPheGlySerTrpThrTyrAspGlyTyrGlnLeuAspLeuGlnLeuGln 191
Db : : : : | | | | : : : :
490 TGCAAACCTGAAGTTTGGTCTCTGCTTACGGAGGCTGGTCTCTGGATCTGCAGATGAG 549

QY 192 AspGluGlyGlyGlyAspIleSerSerPheValThrAsnGlyGluTrpGluLeuGly 211
Db : : : : | | | | : : : :
550 GAG-----GCAGATATCAGTGGCTATATCCCAATGAGAGATGGCACTAGTGGGA 600

QY 212 ValProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIle 231
Db : : : : | | | | : : : :
601 ATCCCCGGCAAGAGAGTGAAGGTTCTATCAGTGTCTGCAAGAGCCCTACCCGATGTC 660

QY 232 ThrPheAlaValValIleArgArgLysThrLeuTyrTyrPhePheAsnLeuIleValPro 251
Db | | | | : : : :
661 ACCTTCAGGTGACATCATCGCGCGAGGACGCTACTATGCGCTCAACTGCTGATCCCC 720

QY 252 CysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlu 271
Db | | | | : : : :
721 TGTGTGCTCATCTCCGCCCTCGCCCTGCTGTGTTCCTGCTTCTCTGAGATTCGCGGGAG 780

QY 272 LysLeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValala 291
Db | | | | : : : :
781 AAGATTTCCTGGGATAACAGTCTTACTCTCTTACCGTCTTATCGCTCTATGCTGCTGGCT 840

QY 292 GluThrMetProAlaThrSerAspAlaValProLeuLeuGlyThrTyrPheAsnCysIle 311
Db : : : : | | | | : : : :
841 GAGATCATGCCCGCAACATCCGATTGCGTACCATGTATGAGCCCAAGTAGTCTGCGGAG 900

QY 312 MetPheMetValAlaSerSerValValSerThrIleLeuIleLeuAsnTyrHisHiArg 331
Db : : : : | | | | : : : :
901 ATGATCATGTGGGCCCTCTCGGTGGTGGTGCAGCGTATGCTGTGCTGATACCCACCAC 960

QY 332 HisAlaAspThrHisGluMetSerAspTrpIleArgCysValPheLeuTyrTrpLeuPro 351
Db | | | | : : : : | | | | : : : :

961 GACCCCGACGGGGCAAGATGCCAAGTGGACAGAGTCACTCTTCTGAACCTGGTGGCG 1020
 QY 352 TrpValLeuArgMetSerArgProGlyL 360
 Db 1021 TGGTTCCTGCGAATGAAGAGCCCGGGGAGGACAAGGTGCCCGGCTGCCAGACAAG 1080
 QY 361 -----SerAlaThrThrProProProAlaArg 369
 Db 1081 CAGCGGGCGCTGACGCTGGCCAGTGTGGAGATGAGCGCCGTGGCGCGCCGCCCGCAGC 1140
 QY 370 -----ValProPro 372
 Db 1141 AACGGGAACCTGCTGTATCATCGGCTTCCGGGGCCTGACGGCGTGCATGTGTCGGGACC 1200
 QY 373 ProProAspLeuGluLeuArgGluArgSerSerLysSerLysSerLeuLeuAlaAsnValLeuAsp 392
 Db 1201 CCCGACTCTGGGTAGTGTGGCGCGCATGGCTGTCTCCCGACGAC----- 1248
 QY 393 IleAspAspAspPheArgHisProGlnAlaGlnProGlnCysCysArgTyrTyrArg 412
 Db 1249 ---GATGAGCACCTCTCTGCAC----- 1266
 QY 413 GlyGlyGluGluAsnGlyAlaGlyLeuAlaAlaHisSerCysPheGlyValAspTyrGlu 432
 Db 1267 GGTGGGCAACCCCCGAGGGG-----GACCCGGAC 1296
 QY 433 LeuSerLeuIleLeuLysGluIleArgValIleThrAspGlnMetArgLysAspAspGlu 452
 Db 1297 TTGGCCAAAGATCTGGAGGAGGTCCGCTACATTGCCAACCGCTTCCGCTGCCAGGACGAA 1356
 QY 453 AspAlaAspIleSerArgAspTyrLysPheAlaAlaMetValValAspArgLeuCysLeu 472
 Db 1357 AGCGAGCGGTCTGCAGCGAGTGGAGTTGCGCCCTGTGTGGTGACCGCTGTGCCCTC 1416
 QY 473 IleIlePheThrLeuPheThrIleIleAlaThrLeuAlaValLeuLeuSerAlaProHis 492
 Db 1417 ATGGCCCTTCGCTCTTCCACCATCTGCACCATCGGCATCTGATGTCGGCTCCCAAC 1476
 QY 493 IleMet 494
 Db 1477 TTCGTG 1482
 RESULT 13
 AAC90380
 ID AAC90380 standard; cdna; 1509 BP.
 XX AC
 AC AAC90380;
 XX DT
 DT 14-MAR-2001 (first entry)
 XX Wild-type human alpha7 ligand gated ion channel coding sequence.
 DE Human; alpha7 nicotinic acetylcholine gated ion channel;
 KW 5-hydroxytryptamine; 5-HT3; calcium ion conductance; ss.
 OS Homo sapiens.
 XX WO200073431-A2.
 PN 07-DEC-2000.
 XX 25-MAY-2000; 2000WO-US011862.
 XX 27-MAY-1999; 99US-0136174P.
 XX (PHAA) PHARMACIA & UPJOHN CO.
 PA Groppi VE, Wolfe ML, Berkenpas MB;
 PI WPI; 2001-061524/P.
 DR P-FSDB; AAB50012.
 XX Special cell culture medium for treating cells and for inducing mammalian


```

QY 352 TrpValLeuArgMetSerArgProGly----- 360
Db 1021 TGGTCTCTGCGATGAAGAGCCCCGGGAGAGAGAGTGGCCCGCTGCCAGCACAAG 1080
QY 361 -----SerAlaThrThrProProAlaArg 369
Db 1081 CAGCGCGCTGCAGCCTGGCAGTGTGGAGATGAGCGCGTGGCGCCCGCCGCCAGC 1140
QY 370 -----ValProPro 372
Db 1141 AACGGGAACCTGCTGTATCATCGGCTTCGCGCCTGGACGGCGTGCACTGTGTCCGACC 1200
QY 373 ProProAspLeuGluLeuArgGluArgSerSerLysSerLeuLeuAlaAsnValLeuAsp 392
Db 1201 CCCGACTCTGGGTAGTGTGGCGCGATGGCCTGCTCCCCACGCAC----- 1248
QY 393 IleAspAspPheArgHisProGlnAlaGlnGlnProGlnCysCysArgTyrTyrArg 412
Db 1249 ---GATGAGCACCTCCTGCAC----- 1266
QY 413 GlyGlyGluGluAsnGlyAlaGlyLeuAlaAlaHisSerCysPheGlyValAspTyrGlu 432
Db 1267 GCGCGGCAACCCCGAGGGG-----GACCCGAC 1296
QY 433 LeuSerLeuLeuLysGluIleArgValIleThrAspGlnMetArgLysAspAspGlu 452
Db 1297 TTGCCCAAGATCCTGGAGGAGTCCGCTACATTGCCAATCGCTTCCGCTGCCAGGACGAA 1356
QY 453 AspAlaAspIleSerArgAspTrpLysPheAlaAlaMetValValAspArgLeuCysLeu 472
Db 1357 AGCGAGCGGTCTGCAGCGAGTGAAGTTGCGCGCTGTGTGGTGACCGCTGTGCTC 1416
QY 473 IleIlePheThrLeuPheThrIleAlaThrLeuAlaValLeuLeuSerAlaProHis 492
Db 1417 ATGCCCTTCGGGTCTTACCATCATCTGCACCATCGGCATCTCTGATGTGGCTCCCAAC 1476
QY 493 IleMet 494
Db 1477 TTCGTG 1482

```

Search completed: May 8, 2004, 19:30:17
Job time : 553.982 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:31:03 ; Search time 12,0702 Seconds
(without alignments)
3952.801 Million cell updates/sec

Title: US-09-303-232-4
Perfect score: 2665
Sequence: 1 MCGRRRRSHLAAPAGLLLL.....LFTIIATLAVLLSAPHIMVS 496

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1267	47.5	502	2 JN0113	nicotinic acetylch
2	1242	46.6	502	2 G02259	alpha 7 neuronal n
3	1241	46.6	502	1 ACHUA7	nicotinic acetylch
4	1241	46.6	502	2 A57175	nicotinic acetylch
5	1232.5	46.2	502	2 T01378	nicotinic receptor
6	1219.5	45.8	511	2 JH0173	alpha-bungarotoxin
7	1201	45.1	498	2 S68588	nicotinic acetylch
8	1141	42.8	461	2 T25671	hypothetical prote
9	1060	39.8	560	2 T19622	hypothetical prote
10	1055.5	39.6	542	2 T19862	hypothetical prote
11	979	36.7	557	2 S12359	nicotinic acetylch
12	962	36.1	503	2 A53956	nicotinic acetylch
13	959	36.0	528	1 ACCH2N	nicotinic acetylch
14	957.5	35.9	567	1 ACPFA1	nicotinic acetylch
15	954	35.8	502	2 A37040	nicotinic acetylch
16	953	35.8	511	2 A40110	nicotinic acetylch
17	936.5	35.1	495	2 S60589	acetylcholine rece
18	932	35.0	499	2 A24572	nicotinic acetylch
19	926	34.7	576	1 ACPFA2	nicotinic acetylch
20	918.5	34.5	456	1 ACCHAN	nicotinic acetylch
21	908.5	34.1	491	1 ACCHNN	nicotinic acetylch
22	904	33.9	622	1 ACCH4N	nicotinic acetylch
23	903.5	33.9	498	2 G02421	nicotinic acetylch
24	902.5	33.9	502	2 S10505	nicotinic acetylch
25	902	33.8	461	2 I50548	nicotinic acetylch
26	901	33.8	503	2 JH0174	acetylcholine rece
27	900.5	33.8	494	2 T03289	nicotinic acetylch
28	899	33.7	461	1 ACRYA1	nicotinic acetylch
29	898.5	33.7	512	2 B37014	nicotinic acetylch

ALIGNMENTS

RESULT 1

JN0113

nicotinic acetylcholine receptor alpha-7 chain precursor, neuronal - chicken
N;Alternate names: alpha-bungarotoxin-binding protein alpha chain
C;Species: Gallus gallus (chicken)

C;Date: 12-Feb-1993 #sequence revision 12-Feb-1993 #text_change 20-Aug-1999
C;Accession: JN0113; JH0172; S28018; B25738; S26566
R;Couturier, S.; Bertrand, D.; Matter, J.M.; Hernandez, M.C.; Bertrand, S.; Millar, N.;

Neuron 5, 847-856, 1990
A;Title: A neuronal nicotinic acetylcholine receptor subunit (alpha 7) is developmental.

A;Reference number: JN0113; MUID:91097796; PMID:1702646

A;Accession: JN0113

A;Molecule type: DNA

A;Residues: 1-502 <COU>

A;Cross-references: GB:X68586; NID:G287756; PIDN:CAA48576.1; PID:G287757

A;Experimental source: white leghorn; brain

R;Schoepfer, R.; Conroy, W.G.; Whiting, P.; Gore, M.; Lindstrom, J.

Neuron 5, 35-48, 1990

A;Title: Brain alpha-bungarotoxin binding protein cDNAs and MABs reveal subtypes of this.

A;Reference number: JH0172; MUID:90315158; PMID:2369519

A;Accession: JN0172

A;Molecule type: mRNA

A;Residues: 1-502 <SCH>

A;Cross-references: EMBL:X52295; NID:G63077; PIDN:CAA36543.1; PID:G63078

A;Experimental source: brain

R;Matter-Sadzinski, L.; Hernandez, M.C.; Rztocil, T.; Ballivet, M.; Matter, J.M.

EMBO J. 11, 4529-4538, 1992

A;Title: Neuronal specificity of the alpha7 nicotinic acetylcholine receptor promoter de

A;Reference number: S28018; MUID:93049204; PMID:1425587

A;Accession: S28018

A;Molecule type: DNA

A;Residues: 1-18 <MAS>

A;Cross-references: EMBL:X68246; GB:S49751; NID:G65319; PIDN:CAA48317.1; PID:G65320

A;Experimental source: white leghorn; erythrocyte

R;Conti-Ironcon, B.M.; Dunn, S.M.J.; Barnard, E.A.; Dolly, J.O.; Lai, F.A.; Ray, N.; Ra

Proc. Natl. Acad. Sci. U.S.A. 82, 5208-5212, 1985

A;Title: Brain and muscle nicotinic acetylcholine receptors are different but homologous

A;Reference number: A94055; MUID:85270494; PMID:3860855

A;Accession: B25738

A;Molecule type: protein

A;Residues: 24-25, 'ET', 28-41, 'X', 43-45, 'X', 47 <CON>

C;Comment: This acetylcholine receptor is blocked by alpha-bungarotoxin and is localized

C;Genetics:

A;Introns: 19/1; 65/3; 80/3; 117/2; 144/1; 200/1; 265/1; 294/1; 330/3

C;Superfamily: acetylcholine receptor

F;1-23/Domain: signal sequence #status predicted <SIG>

F;24-502/Product: nicotinic acetylcholine receptor alpha-7 chain, neuronal #status predi

F;231-254/Domain: transmembrane #status predicted <TR1>

F;262-280/Domain: transmembrane #status predicted <TR2>

F;296-317/Domain: transmembrane #status predicted <TR3>

F;470-488/Domain: transmembrane #status predicted <TR4>

64 DVDEKNQVLTTNIWLQMSWTDHYLQWNVSEYPGVKTVRFPDGGQIWKPKDILLYNSADERFD 123

QY 122 MNSADEGFDSTYPNTNVVRNNGSCLYVPPGIFKSTCKIDITWEPFDDORCEMKFGSWTY 181

Query Match	Score	DB 2;	Length 502;
Post local similarity	46.28	Score 1232.5;	DB 2;
Post local similarity	47.08	Score 1232.5;	DB 2;
Post local similarity	47.08	Score 1232.5;	DB 2;

QY 122 MYNSADEGFDSTYPTNVVVRNNGSCLYVPPGIFKSTCKIDITWFFDDQRCCEMKFGSWTY 181

Db 365 -----NRHSBLIRNIKDNHSLSRANSFPADCRNLQYIMTQSVNSGLTSLSIPSTM 417
QY 424 HSCFGVDYELS-----LIL-----KEIRVITDQMRKDEDEDADISRDKWKFAMVVDRLCLII 474
Db 418 ISSNGTTTDSQATLILHRIYHELVKIVTKRMIEGDEEQACNNKFAAMVVDRLCLIV 477
QY 475 FILETTIATLAVLSAPHIM 494
Db 478 FTIFILVSTIGIFWSAPYIV 497

RESULT 8
T25671
hypothetical protein D2092.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C;Accession: T25671
R;Gattung, S.; Maggi, L.
submitted to the EMBL Data Library, February 1997
A;Description: The sequence of C. elegans cosmid D2092.
A;Reference number: Z20067
A;Accession: T25671
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-461 <GAT>
A;Cross-references: EMBL:U98167; PIDN:ABA42223.1; GSPDB:GN00019; CESP:D2092.3
A;Experimental source: strain Bristol N2; clone D2092
C;Genetics:
A;Gene: CESP:D2092.3
A;Map position: 1
A;Introns: 36/1; 80/3; 119/2; 191/1; 243/1; 274/1; 303/1; 382/2
C;Superfamily: acetylcholine receptor

Query Match 42.8%; Score 1141; DB 2; Length 461;
Best Local Similarity 46.2%; Pred. No. 1.5e-91;
Matches 218; Conservative 85; Mismatches 121; Indels 48; Gaps 7;

QY 34 EKRLHLLHDYNNVLRPVNNSDPLQSLFGLTLMQIIDVDKKNOLLINWIKLWNDM 93
Db 26 ETKLFDTLLKGNYPLEPQVNSQPLEVKIKLFQQLIDVDKKNVSNVNWLSYTFDH 85

QY 94 NLRWNTSDFGVKDLRVP--PHRLWKPDLVLYNSADEGDSYPTVTVNVRNNSCLYVPP 151
Db 86 KLOWEPKYYGIGQIDIFPGSSDHINWKPDLVLYNSADEGSTFKSNLLTHTGTVWIPP 145

QY 152 GIFKSTCKIDITWPPDDQRCWKFGSWYDYGQLDLQDEGG-----DISSFVNGEW 207
Db 146 GVLKFCVQLDVTWFPEDQVCEMKFGSWTFHGYAIDLIQIDDDTNGTQSMDLSTLVNGEW 205

QY 208 ELIGVPGKNEIYYNCCPEYIDITFAVVRIRKTLAYFFENLIVPCVLIASMALLGFTLP 267
Db 206 QVLTWAKRVSYKCCPEYPTVNYLHRRRTLYGFENLILPSLLISLMALIGWFPP 265

QY 268 DSGEKLGLVITLLSLTVFLNVAETPATSADVLLGTYNCFIMFWVASSVSTILIN 327
Db 266 DAGEKITLVILLAVFLSMVSEMTPTTSEAVPLIGVFFSCMLVSVSVFTIVLN 325

QY 328 YHHRHADTHMSDWIRCVFLYWLPMWLMSRPSGATPTPPARVPPPPDLELRSSKLL 387
Db 326 LHFRSADSHENPLVRVLEFLPMLFMSRPG-----YKFKV 363

QY 388 ANVLIDDDPHRPAQOQPCRYRGEENGAGLAHSCFGVDYELSLII-----KEIRV 442
Db 364 ANVIDSTDKM-PKKKNPLDCNL-----PSNHAG-----YEAQILLHSHVHTELRR 408

QY 443 ITQMRKDDSDADISRDKWKFAMVVDRLCLIIIFTLTIIATLAVLSAPHIM 494
Db 409 VVAFYNKEEHEDEIQTDRFAMVVDRLCLIIITFTVIVISILAMWSAPHII 460

RESULT 9
T19622
hypothetical protein C31H5.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T19622
R;Hembry, C.
submitted to the EMBL Data Library, March 1996
A;Reference number: Z19188
A;Accession: T19622

C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C;Accession: T19622
R;Kershaw, J.
submitted to the EMBL Data Library, April 1997
A;Reference number: Z19153
A;Accession: T19622
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-560 <WIL>
A;Cross-references: EMBL:Z93778; PIDN:CAB07843.1; GSPDB:GN00019; CESP:C31H5.3
A;Experimental source: clone C31H5
C;Genetics:
A;Gene: CESP:C31H5.3
A;Map position: 1
A;Introns: 24/1; 70/3; 139/2; 270/1; 299/1; 336/3; 372/2; 456/3
C;Superfamily: acetylcholine receptor

Query Match 39.8%; Score 1060; DB 2; Length 560;
Best Local Similarity 37.6%; Pred. No. 2.4e-84;
Matches 217; Conservative 96; Mismatches 142; Indels 122; Gaps 11;

QY 16 LMLLLCLMPRGARC-----GYHEKRLHLLHDYNNVLRPVNNSDPLQSLFGLTLMQ 69
Db 9 VLSVSLIWE--TKCSKVIWTGDEHERLYAKLAENYNKLARPVNNESEAVVLLGMDYQ 66

QY 70 IIDVDKKNOLLINWIKLWNDMNRWNTSDFGVKDLRVPVPHRLWKPDLVLYNSADEG 129
Db 67 ILDIDEKHQIMNSVNLWRNSWTDHYLTWDPSEFGNIKEVRLPINNWKPDVLLYNSVDQ 126

QY 130 FSTYPTNVVVRNNSCLVVPGIFKSTCKIDITWPPDDQRCWKFGSWYDYGQLDLQ 189
Db 127 FSTVPMNAVLYTGNVTWIPPAIRSSCAIDIAFFDTQHCTMFGSWYSGFTDL- 185

QY 190 LODEGGDISFVT-----NGEWELIGVPGKNEIYYNCCPEYIDITFAVVRIRKTL 243
Db 186 -----INTVISPATYKPNGEWELLGTSQRSIFFECCPEYDYVTFVSIIRRTLY 237

QY 244 YFNLIIVCVLIASMALLGFTLPDSEKLSLGVITLLSLTVFLNVAETMPTSDAVPL 303
Db 238 YGFNLIPLCWLISLALLSFTLPADCGEKLNLGVTFMSLCVFMIMVAEAMPOTSALPL 297

QY 304 LGTYFNCIMFWVASSVSTILINVYHRHADTHE-MSDWIRCVFLYWLPMWLMSRPSGA 362
Db 298 IQTYFCIMFWQVGSVAIVIALNFHRSPEQTKPMKNKFLKLLGLLGLWLTLLGWERPDV- 356

QY 363 TTPPPARVPPPDLEL-----RERSKSLANV-----L 391
Db 357 -----LELSVGHAYASDNKKQYLYEVERHILTRPENGHGSADVKAHL 403

QY 392 DIDDDPHRPAQOQ-----CCYYRGGGEENGA 419
Db 404 DLSTGNPHSDAKKSPSPKRTSASIMGMTGLPTOMNGALDSSINKYCTKTVTRPLENGS 463

QY 420 GLAAH-----SCFG-----VDYELSLILKEIRVITDQMRKDEDEDISR 458
Db 464 ATINHKSSQIPINNNTYKANNOKTQFDRHFHILNELRVISARVKEEMHALQA 523

QY 459 DWKFAAMVVDRLCLIIIFTLTIIATLAVLSAPHIMV 495
Db 524 DMWFASRVVDRVCFLAFSAFLFMCITAIISYNAPHLFV 560

RESULT 10
T19862
hypothetical protein C40C9.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T19862
R;Hembry, C.
submitted to the EMBL Data Library, March 1996
A;Reference number: Z19188
A;Accession: T19862

A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-542 <WIL>
A;Cross-references: EMBL:Z70266; PIDN:CAA94206.1; GSPDB:GNO0028; CESP:C40C9.2
A;Experimental source: clone C40C9
C;Genetics:
A;Gene: CESP:C40C9.2
A;Map position: X
A;Insertions: 11/3; 69/3; 123/2; 173/3; 216/1; 248/1; 279/1; 314/1; 350/3; 430/1; 464/1; 49
C;Superfamily: acetylcholine receptor

Query Match 39.6%; Score 1055.5; DB 2; Length 542;
Best Local Similarity 40.5%; Pred. No. 5.6e-84;
Matches 214; Conservative 94; Mismatches 136; Indels 85; Gaps 12;

QY 34 EKRLHLLHLLHYNLRLPVPVNESDPLQSLFGLTLMQIIDVDKKNLLITNWLKLENDM 93
DB 30 EYRLADLRNLDYERPVANASEPLVSVVKIYLQQILDVDKKNQVITLVAMISYQWTDY 89
QY 94 NLRWNTSDFGVKDLRVP--PHRLWKPDVLMYNSADEGFDSTYPTNVVVRNNGSLYVPP 151
DB 90 KKWDPSEYGGIKDIRIPGNANALWKPDVLLYNSADENFDSTYPTNVVSVYTGDLQVPP 149
QY 152 GIFKSTCKIDITWFPDDQRCCKMFGSWTGYQLDQLQ-----DEGGDISSVFTN 204
DB 150 GILKLSCKIDITYPPDDQICHLKFGSWTYSGNFIDLRINGPEGKNISDEGIDVQYVQV 209
QY 205 GEMELIGVGRNEIYVNCPPERYDITFAVIRRKTLVYFFENLIVPCVLIASMALLGET 264
DB 210 GEWNLAVPARHETNFD--EQPSPFFYLIQRTYLYGLNLIISFLISLMTVLGET 267
QY 265 LPDPSGKLSGLVTILSLVFLMVAETMPATSDAVPELL-----GTYFNCIMFVASS 318
DB 268 LPDAGEKITLEITILLSVCFELSMVADMTPTSEAVPLIGLIIFSGAFTSCCMLVVSAS 327
QY 319 VVSTILILNHYHRHADTHEMSDWIRCVFLYWL PWLVRMSRPG-----SATTPPPA 368
DB 328 WVFTVLNHLNHRKPETHMSPEFLRELLIWL PWLLLMRRPFGKTFNCTHLKAEKAEKA 387
QY 369 RV-----PPPDLELRSSKSLANV-----LDIDDDF--RH-----PQAOQ 404
DB 388 KQSIKNGVCGKPTSDVHPSGSLMKNTKLGQQIIDFEYEHVQHNHLMFVAPSEMT 447
QY 405 PQCC-----RYRGGENGAGLAHSCFGVDYELSLILKEIRVITD 445
DB 448 PRVYSKWMAESYVEDVMTLNKMQ-----KACL-----ELKNISSQTRAMRK 492
QY 446 QMRKDEDEDADISRDWKFAMVVDRLCLIIITLTIATLAVILSAPHIM 494
DB 493 KMBEDERQAAQNDWKFAMVVDRLCLITFSVFIWSTGIMFSSPHLI 541

RESULT 11
S12359
nicotinic acetylcholine receptor alpha-L1 chain precursor - desert locust
C;Species: Schistocerca gregaria (desert locust)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Aug-1999
C;Accession: S12359
R;Marshall, J.; Buckingham, S.D.; Shingai, R.; Lunt, G.G.; Goosey, M.W.; Darlison, M.G.;
EMBO J. 9, 4391-4398, 1990
A;Title: Sequence and functional expression of a single alpha subunit of an insect nicot
A;Reference number: S12359; MUID:91092263; PMID:1702381
A;Accession: S12359
A;Molecule type: mRNA
A;Residues: 1-557 <MAR>
A;Cross-references: EMBL:X55439; NID:g10133; PIDN:CAA939081.1; PID:g10134
C;Superfamily: acetylcholine receptor
C;Keywords: Glycoprotein; Ion channel; neurotransmitter receptor; transmembrane protein
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-557/Product: nicotinic acetylcholine receptor alpha-L1 chain #status predicted <MAT
F;245-266/Domain: transmembrane #status predicted <TM1>
F;274-295/Domain: transmembrane #status predicted <TM2>
F;308-329/Domain: transmembrane #status predicted <TM3>

F;501-523/Domain: transmembrane #status predicted <TM4>
F;47,235/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 36.7%; Score 979; DB 2; Length 557;
Best Local Similarity 40.0%; Pred. No. 2.8e-77;
Matches 217; Conservative 88; Mismatches 160; Indels 78; Gaps 10;

QY 11 AAPAGLILLCLLWPGRCARCYHEKRLHLLHLLHYNLRLPVPVNESDPLQSLFGLTLMQI 70
DB 4 ALPEMLILLLLHHPAAANPDARLYDDLLSYNRLIRPVSNNTTVLVKGLRLSQL 63
QY 71 IDVDEKNQLITNWLKLEWMDNMLRWNTSDFGVKDLRVPVPPHPLKPDVLMYNSADEGF 130
DB 64 IDLMLKQILTTNWLSEWQDHKFRWDPAEYGGVTELYVSEHILWLPDILVYNNADGEY 123
QY 131 DSYPTNVVVRNNGSLYVPPGIFKSTCKIDITWFPDDQRCCKMFGSWTGYQLD-- 188
DB 124 VVTMTKAVLHHTGKVVMTTPAIFKSSCEIDVRYFPDQQTCFMKFGSWTGYDQIDLKH 183
QY 189 --QLQDEG----GGDISSVFTNGEWELIGVGRNEIYVNCPPERYDITFAVIRRKTL 242
DB 184 INQYDNDKVKVGLDIREYVPSVENDILGVPAERHEKYPCCAEPYDIFENITLRKTL 243
QY 243 YFFNLIVPCVLIASMALLQFTLPDPSGKLSGLVTILLSLTVFLNMVAETMPATSDAVP 302
DB 244 FYTNLIVPCVIGISYLSVLVFLPADSGEKIALCISILLQTMFFLLISEIIPSTSLAP 303
QY 303 LLGYTFNCIMFVASSVSVTILNHYHRHADTHEMSDWIRCVFLYWL PWLVRMSRPGSA 362
DB 304 LLGKYLFTVVLVGLSVVITIMVNLNHYRKPSTHKMAPVWRKVIRLPLKLLM----- 357
QY 363 TTPPPARVPPPDLELRSSKSL-----ANVLIDDDDFRHPQ 401
DB 358 -----RV---PQQLADLASKRLHRAHNSKLSAAAAAASASSPSDLRHH 408
QY 402 AQDPQCCRY-----YRGGENGAGLAH-----SCFGVD----- 430
DB 409 LHQHQHQLHQLHHLQRPQCGNGLHSATNFGSGAGAFGLPSVVGDLGSLSDVATRKKY 468
QY 431 -YELSLILKEIRVITDQMRKDE-DADISRDWKFAMVVDRLCLIIITLTIATLAVIL 488
DB 469 PFELEKAHNVLFIQNHMQRQDFDAB-DQDWGFVAMVLDRLFLWIFIETIASIVGTFAIL 527
QY 489 SAP 491
DB 528 EAP 530

RESULT 12
A53956
nicotinic acetylcholine receptor alpha-3 chain precursor, neuronal - human
C;Species: Homo sapiens (man)
C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 20-Aug-1999
C;Accession: A53956; S21338
R;Minovilovic, M.; Roses, A.D.
Exp. Neurol. 111, 175-180, 1991
A;Title: Expression of mRNAs in human thymus coding for the alpha3 subunit of a neuronal
A;Reference number: A53956; MUID:91114756; PMID:1989896
A;Accession: A53956
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-503 <MIH>
A;Cross-references: GB:M37981; NID:g189252; PIDN:AAA59942.1; PID:g189253
R;Anand, R.; Lindstrom, J.
submitted to the EMBL data Library, June 1990
A;Description: Nucleotide sequence of the mature human nicotinic acetylcholine receptor
A;Reference number: S21338
A;Accession: S21338
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 30-503 <ANA>
A;Cross-references: EMBL:X53559; NID:g34985; PIDN:CAA37625.1; PID:g34986
C;Genetics:

A:Gene: GDB:CHRNA3
A:Cross-references: GDB:125219; OMIM:118503
A:Map position: 15q24-15q24
C:Superfamily: acetylcholine receptor
C:Keywords: neurotransmitter receptor

Query Match 36.1%; Score 962; DB 2; Length 503;
Best Local Similarity 41.1%; Pred. No. 7.3e-76;
Matches 206; Conservative 89; Mismatches 172; Indels 34; Gaps 11;
QY 10 LAAPAG---LILLICLLPRGARGCYHEKRLHLLHDHYNVLERPVNNSDPLQLSFGL 65
DB 6 LAAPGAVARLLLLLSLIP- VARASEAERLPERLFEDYNEIRPVANVSDPVIIHFV 64
QY 66 TLNQIIDVDEKQQLLTINLWLKLEWDMNLRWNTSDFGVKDLRVPPHRLMKDPDVLMYNS 125
DB 65 SMSQLVKVDEVNQIMETNLWLKQIMNDYKLVKWPSPDYGGAERMRVPAQKIWKPDIVLYNN 124
QY 126 ADEGFDSTYPTNVVVRNNGSLVPPGIKSTCKIDITWFFPDORCEMKFGSWTYDGYQ 185
DB 125 AVGDFOVDKTKALLKYTGVTWIPPAIFKSSCKIDITVPPFDYQNTKFKGWSYDKAK 184
QY 186 LDQLQDEGGDISSTFWNGEWELIGVPGKRNEIYVNCPEYIDITFAVIRRKTLTYF 245
DB 185 IDLVIG-SSMNLKDYWSEGEWALIKAPGYKHIDKNCCEIYPDITYSLYIRRLPLYT 243
QY 246 FNLIVPCVLIIASNALIGFTLPDPSGKLSGVTLISLTVFLNMVAETMPATSDAVPLLG 305
DB 244 INLIIPCLLISFLTVLVFVLPSCGKVLICISVLLSLTVFLVITETIPSTSLVPLIG 303
QY 306 TYPNCITFMVAVSSVSTILINLHHRHADHENSMDIRVCVLYWLPWLMRSP- - - - -GS 361
DB 304 EYLLTFMIFVTLISIVTVFLNVHYRTPTHTWPSWKTVFLNLLPRVPMFTPTNEGN 363
QY 362 ATTPPARVPPPPPLELRS- SKSLANLVLDIDDDFRHPQAQPOC- - - - -CRYRGGBEN- 417
DB 364 AQPRFLYGAELSNLCFSAESKGKEG- - - - -YP-CDQMGYCHHRIKISNF 413
QY 418 GAGLA-AHSCFGVDYELSL- - - - -ILKEIRVITDMQRKDDADISRDKFAAMVV 467
DB 414 SANLTRSSSESVDVLSLSALSPETKEAIOQVKYIAENKMAQNEAKEIQDDWKYVAMVI 473
QY 468 DRCLLIIFTLFTIATLAVLL 488
DB 474 DRIFLWVFTLVCLITAGLFL 494

RESULT 13

ACCH2N

nicotinic acetylcholine receptor alpha-2 chain precursor, neuronal - chicken
C:Species: Gallus gallus (chicken)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Jul-2000
C:Accession: S00377
R:Nef, P.; Oneyser, C.; Alliod, C.; Couturier, S.; Ballivet, M.
EMBO J. 7, 595-601, 1988
A:Title: Genes expressed in the brain define three distinct neuronal nicotinic acetylcholine receptor subunits
A:Reference number: S00376; MUID:88283624; PMID:3267226
A:Accession: S00377
A:Molecule type: DNA
A:Residues: 1-528 <NEF>
A:Cross-references: EMBL:X07339; NID:g62792; PIDN:CAB59645.1; PID:g6136914
C:Genetics:
A:Introns: 21/1; 73/3; 88/3; 125/2; 487/3
C:Superfamily: acetylcholine receptor
C:Keywords: glycoprotein; ion channel; neurotransmitter receptor; postsynaptic membrane;
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-528/Product: nicotinic acetylcholine receptor alpha-2 chain #status predicted <MAT>
F:241-264/Domain: transmembrane #status predicted <TM1>
F:272-290/Domain: transmembrane #status predicted <TM2>
F:306-327/Domain: transmembrane #status predicted <TM3>
F:502-520/Domain: transmembrane #status predicted <TM4>
F:54, 104/Binding site: carboxylate (Asn) (covalent) #status predicted
F:158-172, 222-223/Disulfide bonds: #status predicted

Query Match 36.0%; Score 959; DB 1; Length 528;
Best Local Similarity 39.2%; Pred. No. 1.4e-75;
Matches 198; Conservative 84; Mismatches 161; Indels 62; Gaps 10;

QY 31 GYHEKRLHLLHDHYNVLERPVNNSDPLQLSFGLTLMQIIDVDEKQQLLTINLWIKLEW 90
DB 31 GFAERLFLKHLFTGYNRWSRFPVNTSDVIVFGLSIAQLIDVDEKQMTINWLKQEW 90
QY 91 NDMNLRWNTSDFGVKDLRVPPHRLMKDPVLMYNSADEGFDSTYPTNVVVRNNGSLVYP 150
DB 91 SDYKLRWNPEDFDNVTISIRVSEMIWIPDILVYNNADGEFAVTHMTKAHLFSNGKVKWP 150
QY 151 PGIFKSTCKIDITWFFPDORCEMKFGSWTYDGYQLDLQLODEGGDISSTFWNGEWELI 210
DB 151 PAIYKSSCSIDVTPFPDQCNCKMKFGSWTYDKAKIDLE-NMEHHVDLKDYESEGEWAI 209
QY 211 GYVGRKNEIYVNCPEYIDITFAVIRRKTLTYFPNLIYVPCVLIIASMALIGFTLPDPSG 270
DB 210 NAIGRYSKKYDCCIEIYDITFYVIRRLPLYFTINLIIPCLLISLTVLVFVLPDSCG 269
QY 271 EKLSLGVTLISLTVFLNMVAETMPATSDAVPLLGTYFNCIMPMVAVSSVSTILINLHHR 330
DB 270 EKITLCISVLLSLTVFLVITETIPSTSLVPLIGSYLLFTMIFVTLISITVFLNVH 329
QY 331 RHADTHEMSDWIRVCVLYWLPWLMRSPGSAATPPPARVP- - - - -PPDLELRERS 382
DB 330 RSPSTHMPHWRSFPLGFIPRLFMKR- - - - -PPLLLPAEGTTGQVDPGTRL- - - - -ST 380
QY 383 SKSLANLVLDIDDDFRHPQAQPOC- - - - -CCRYRGGEENGAGLAH-SC- - - - -426
DB 381 SRCWLET- - - - -DVDDKWEEREEEEEEEEEEKAYSRVPSGSG- - - - -GTQCHYSERQA 435
QY 427 - - - - -FGVDYELSL- - - - -ILKEIRVITDMQRKDDADISRDKFA 463
DB 436 GKASGGPAQVPLKGEVSGDQGLTSLPSILRALEGVQVIADHLRAEDADFSVKEDWKYV 495
QY 464 AMVVDRLCLLIIFTLFTIATLAVLL 488
DB 496 AMVIDRIFLWMFTIVCLITAGLFL 520

RESULT 14

ACFPAL

nicotinic acetylcholine receptor alpha-1 chain precursor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 21-Jun-2002
C:Accession: S00381; A38801
R:Bossey, B.; Ballivet, M.; Spierer, P.
EMBO J. 7, 611-618, 1988
A:Title: Conservation of neural nicotinic acetylcholine receptors from Drosophila to vertebrates
A:Reference number: S00381; MUID:88283626; PMID:2840281
A:Accession: S00381
A:Molecule type: DNA
A:Residues: 1-567 <BOS>
A:Cross-references: GB:X07194; NID:g7575; PIDN:CAA30172.1; PID:g7576
A:Accession: A38801
A:Molecule type: mRNA
A:Residues: 1-567 <BOS>
A:Cross-references: EMBL:X07194; NID:g7575; PIDN:CAA30172.1; PID:g7576
A:Note: 538-Tyr was also found
C:Genetics:
A:Gene: FlyBase:ncr-alpha-96Aa
A:Cross-references: FlyBase:FBgn0000036
A:Map position: 3R 96A
A:Introns: 64/3; 79/3; 116/2; 176/3; 330/2; 401/1; 499/3
C:Superfamily: acetylcholine receptor
C:Keywords: glycoprotein; ion channel; neurotransmitter receptor; postsynaptic membrane
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-567/Product: nicotinic acetylcholine receptor alpha-like chain #status predicted <F>
F:22-240/Domain: extracellular #status predicted <EXT>
F:240-264/Domain: transmembrane #status predicted <TM1>
F:272-290/Domain: transmembrane #status predicted <TM2>

F:306-325/Domain: transmembrane #status predicted <TM3>
F:326-513/Domain: intracellular #status predicted <INT>
F:514-532/Domain: transmembrane #status predicted <TM4>
F:45,233/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:149-163,222-223/Disulfide bonds: #status predicted

Query Match 35.8%; Score 957.5; DB 1; Length 567;
Best Local Similarity 38.2%; Pred. No. 2.1e-75;
Matches 206; Conservative 86; Mismatches 174; Indels 73; Gaps 11;

```
QY 18 LLLCLLPRGARGCYHEKRLHLLDHYNVLRVAVNESDPLQLSFGTLMOITDVEKN 77
Db 9 VFIALHATGGLANDAKRKYDDLLSNYRLIRVGNNSDRLTVMKGLRLSQLDVLNKN 68
QY 78 QLLLTINWLEMDNMLRWNTSDFGVKDLRVPPLHLPKDPDLVLMYNSADGSPSTPTN 137
Db 69 QIMTNVVEGWNDYKLNWPDYGGVDTLHVSEHLWHPDVIYNNADGNYEVTIMTK 128
QY 138 VVRNNGSLVPPGIFKSTCKIDI TWFPDDQCEMKFGSWTYDGYOLDL----QLQD- 192
Db 129 AILHHTGKVMKPPAIYKSFCEIDVEYFPFDEQTCFMKFGSWTYDGYVLDLRLKQTADS 188
QY 193 ---EGGDDISSFTVNGEWELIGVPGKNEIYNNCCPEYIDITFAVIRKTLYYFNL 249
Db 189 DNEVIGDLQDYIISVEWDMRVPVRNEKPYSCCEPYLDIVENLTIRKTLFYTNLI 248
QY 250 VPCVLIA SMALLGFTLPDSEKLSGLVTILLSTVFLNMVAETMPATSDAVPLLGTYFN 309
Db 249 IPCVGISFLSVLFVLPDSGEKLSLCISILSLTVFLLAEIIPTSLSVPLLGKYL 308
QY 310 CIMFWASSVSTILLIYHHRHADTHEMSDWIRCVFLYWL PWVLRMSGSAATPPPAR 369
Db 309 FTMMLVTLVVVTIAVLNVNFRSPVTHRMAPVQRLFIQILPKLCLIERP--KKEEPEED 366
QY 370 VPP-----PPDL-LRERSKSL-----L 387
Db 367 QPEVLTDVHLPPDDVFNYSKRFSGDYGPALPASHRFDIAAGGISAHCFAEPPL 426
QY 388 ANVLDI---DDDFRHPOAQ----QPCCRYRGGGEAGLAHS-----CFGVDY--ELS 434
Db 427 PSSLPFGADDDLFSPLGNDISPGCC-----PAAAAAADA LSPTEKPYAREME 478
QY 435 LIKEIRVITDQMKDEDEDADISRDWKFAMVDRLCLITFTTIATLAVLSAPHI 493
Db 479 KTEGSRFIAQHVKNKDKFESVEDWKYVAMVLDRMFLWIFAIACVVGVTALIIQADSL 537
```

RESULT 15

A37040
nicotinic acetylcholine receptor alpha-3 chain precursor, neuronal (version 2) - human
C/Species: Homo sapiens (man)
C/Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 21-Jul-2000
C/Accession: A37040; S24595
R/Fornasari, D.; Chini, B.; Tarroni, P.; Clementi, F.
Neurosci. Lett. 111, 351-356, 1990
A/Title: Molecular cloning of human neuronal nicotinic receptor alpha-3-subunit.
A/Reference number: A37040; MUID:90245296; PMID:2336208
A/Accession: A37040
A/Molecule type: mRNA
A/Residues: 1-502 <FOR>
A/Cross-references: EMBL:X52239; NID:g177897; PIDN:AAC84176.1; PID:g177899
C/Superfamily: acetylcholine receptor
C/Keywords: neurotransmitter receptor; transmembrane protein
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-502/Product: nicotinic acetylcholine receptor alpha-3 chain #status predicted <MAT>

Query Match 35.8%; Score 954; DB 2; Length 502;
Best Local Similarity 40.8%; Pred. No. 3.6e-75;
Matches 203; Conservative 90; Mismatches 174; Indels 30; Gaps 10;

```
QY 10 LAAPAGLLLCLLMPRGARGCYHEKRLHLLDHYNVLRVAVNESDPLQLSFGTLMOQ 69
Db 9 LACRARILLLLLSLP-VARASEAERHLPFLFEDYNEIRPVANVSDPVIHFEVMSQ 67
```

```
QY 70 IIDVDEKNQLITINWLEMDNMLRWNTSDFGVKDLRVPPLHLPKDPDLVLMYNSADEG 129
Db 68 LVKDEVNQIMETINWLEKQIWNQYKLNWPSGYGAEFMRVPAQKIWKPDVILVNNVAVGD 127
QY 130 FDSYPTNVVVRNNGSLVPPGIFKSTCKIDI TWFPDDQCEMKFGSWTYDGYQLDLQ 189
Db 128 FOVTTKTKALLKYTGVTWIPPAIFKSSCKLDVYFFDYQNCCTMKFGSWSYDKAKIDL 187
QY 190 LQDEGGDISFTVNGEWELIGVPGKNEIYNNCCPEYIDITFAVIRKTLYYFNL 249
Db 188 LIG-SSMWLKDYSWESGEWAIKAPGYKHDIKYNCCEIYPDITYSLYSRRRLPFLV 246
QY 250 VPCVLIA SMALLGFTLPDSEKLSGLVTILLSTVFLNMVAETMPATSDAVPLLGTYFN 309
Db 247 IPCLLISFLTVLFVLPDSGEKLSGLVTILLSTVFLNLLTPSTLSVPLIGEVLL 306
QY 310 CIMFWASSVSTILLIYHHRHADTHEMSDWIRCVFLYWL PWVLRMSRSP---GSATTP 365
Db 307 FTMIFVTLIVITVFLNVHVRTPPTHTMPSWVKTVFLNLLPRVMEWTRPTSNENAKQP 366
QY 366 PPARVPPPPDLRLERS-SKSLIANVLDDDDDFRHPOAQOPOC--CRYRGGGEEN-GAGL 421
Db 367 RPLYGAELSNLNCFSRAESKCKEG-----YP-CQDGMCGYCHHRIKISNFSANL 416
QY 422 A-AHSCFGVDYELSL-----ILKEIRVITDQMKDEDEDADISRDWKFAMVVDRLC 471
Db 417 TRSSSESVDVAVLSLSALSPEIKEAIQSVKYIAENMKAQNEAKEIQDDWKYVAVIDRIF 476
QY 472 LIIFTFTIATLAVLL 488
Db 477 LWVFTLVCIILGTAGLFL 493
```

Search completed: May 7, 2004, 11:39:31
Job time : 13.0702 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:18:28 ; Search time 9.26316 Seconds
(without alignments)
2788.120 Million cell updates/sec

Title: US-09-303-232-4
Perfect score: 2665
Sequence: 1 MGRARRSHLAAPAGALLLL.....LFTIITLAVLLSAPHIMVS 496

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1267	47.5	502	1	ACH7_CHICK
2	1246	46.8	502	1	ACH7_HUMAN
3	1241	46.6	502	1	ACH7_MOUSE
4	1232.5	46.2	502	1	ACH7_RAT
5	1218.5	45.7	499	1	ACH7_BOVIN
6	1201	45.1	498	1	ACH1_CAEEL
7	980.5	36.8	516	1	ACH1_MANSE
8	979	36.7	557	1	ACH1_SCHGR
9	960	36.0	503	1	ACH3_HUMAN
10	960	36.0	529	1	ACH2_HUMAN
11	959	36.0	528	1	ACH2_CHICK
12	956.5	35.9	567	1	ACH1_DROME
13	951	35.7	511	1	ACH2_RAT
14	936.5	35.1	495	1	ACH3_BOVIN
15	930	34.9	499	1	ACH3_RAT
16	926	34.7	576	1	ACH2_DROME
17	918.5	34.5	456	1	ACHA_CHICK
18	918	34.4	434	1	ACH6_HUMAN
19	917	34.4	496	1	ACH3_CHICK
20	908.5	34.1	491	1	ACHN_CHICK
21	906	34.0	461	1	ACHA_TORMA
22	904	33.9	622	1	ACHA_CHICK
23	903.5	33.9	498	1	ACHP_HUMAN
24	902.5	33.9	502	1	ACHN_HUMAN
25	902	33.8	495	1	ACHP_RAT
26	899	33.7	461	1	ACHA_TORCA
27	898.5	33.7	512	1	ACH3_CARAU
28	898	33.7	500	1	ACHN_RAT
29	896.5	33.6	457	1	ACHA_BOVIN
30	896.5	33.6	493	1	ACH6_RAT
31	893.5	33.5	519	1	ACH4_DROME
32	890	33.4	457	1	ACHA_RAT
33	889	33.4	457	1	ACHA_MOUSE

34	888	33.3	470	1	ACHP_CHICK
35	886.5	33.3	494	1	ACH6_CHICK
36	885.5	33.2	459	1	ACHN_CARAU
37	880	33.0	456	1	ACHA_BRARE
38	879.5	33.0	627	1	ACH4_HUMAN
39	876.5	32.9	457	1	ACH2_XENLA
40	874	32.8	521	1	ACH3_DROME
41	872	32.7	538	1	ACH8_CAEEL
42	867.5	32.6	630	1	ACH4_RAT
43	866.5	32.5	482	1	ACHA_HUMAN
44	865.5	32.5	629	1	ACH4_MOUSE
45	858.5	32.2	452	1	ACH5_RAT

ALIGNMENTS

RESULT 1
ACH7_CHICK
ID ACH7_CHICK STANDARD; PRT; 502 AA.
AC P22770;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-7 chain precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=91097796; PubMed=1702646;
RA Couturier S., Bertrand D., Matter J.M., Hernandez M.C., Bertrand S.,
RA Millar N., Valera S., Barkas T., Ballivet M.;
RT "A neuronal nicotinic acetylcholine receptor subunit (alpha 7) is
RT developmentally regulated and forms a homo-oligomeric channel blocked
RT by alpha-BTX.";
RL Neuron 5:847-856(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=90315158; PubMed=2369519;
RA Schoepfer R., Conroy W.G., Whiting P., Gore M., Lindstroem J.;
RT "Brain alpha-bungarotoxin binding protein cDNAs and MABs reveal
RT subtypes of this branch of the ligand-gated ion channel gene
RT superfamily";
RL Neuron 5:335-48(1990).
RN [3]
RP SEQUENCE OF 1-18 FROM N.A.
RC STRAIN=White leghorn; TISSUE=Erythrocyte;
RX MEDLINE=93049204; PubMed=1425587;
RA Matter-Sadzinski L., Hernandez M.C., Roztocil T., Ballivet M.,
RA Matter J.M.;
RT "Neuronal specificity of the alpha 7 nicotinic acetylcholine receptor
RT promoter develops during morphogenesis of the central nervous
RT system.";
RL EMBO J. 11:4529-4538(1992).
RN [4]
RP SEQUENCE OF 24-47.
RC TISSUE=Brain;
RX MEDLINE=85270494; PubMed=3860855;
RA Conti-Tronconi B.M., Dunn S.M.J., Barnard E.A., Dolly J.O., Lai F.A.,
RA Ray N., Rafferty M.A.;
RT "Brain and muscle nicotinic acetylcholine receptors are different but
RT homologous proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:5208-5212(1985).
RN [5]
RP MUTAGENESIS OF LEU-270.
RX MEDLINE=92049732; PubMed=1719423;
RA Recah F., Bertrand D., Galzi J.-L., Devillers-Thierry A., Mulle C.,
RA Hussy N., Bertrand S., Ballivet M., Changeux J.-P.;

QY 312 MEMVASSVSTLIILNHHRRHADTHEMSDWIRCVFLYMLPWLRLMRSPG----- 360
 Db 301 MIIVGLSVVTVIVLQYHHHDPDGGKMPKWTIRVILLNCAWFLMRKRGEDKVRPACQHK 360
 QY 361 -----SATTTPPAR-----VPPDLELRERSSKSLIANVLD 392
 Db 361 QRCSLASVEMGAVAPPASNGNLLYIGRGLDGHCVFTPDGSGVVCGRMACSPH----- 416
 QY 393 IDDDFRHQAQPPCCRYRGGEAGLAHSCFGVDYELSILKEIRVITDQMKODE 452
 Db 417 -DEHLH-----GGQPPG-----DPDLAKILEVRYIANRFRCODE 452
 QY 453 DADISDDWFAAMVVDRLCLIIFTLTIIATLAVLSAPHIM 494
 Db 453 SEAVCSEWFAACVVDRLCLMAFSVFTIICTIGILMSAPNFV 494

RESULT 3
 ACH7_MOUSE STANDARD; PRT; 502 AA.
 AC P49582;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neuronal acetylcholine receptor protein, alpha-7 chain precursor.
 GN CHRNA7 OR ACRA7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Brain;
 RA MEDLINE=95324936; PubMed=7601470;
 RX Orr-Urtreger A., Seldin M.F., Baldini A., Beaudet A.L.;
 RT "Cloning and mapping of the mouse alpha 7-neuronal nicotinic
 acetylcholine receptor.";
 RL Genomics 26:399-402(1995).
 CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
 extensive change in conformation that affects all subunits and
 leads to opening of an ion-conducting channel across the plasma
 membrane.
 CC -!- SUBUNIT: Forms a homo-oligomeric channel blocked by alpha-
 bungarotoxin. The structure is probably pentameric (5y
 similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L37663; AAC42053.1; -.
 DR PIR; A57175; A57175.
 DR MGI; MGI:99779; Chrna7.
 DR InterPro; IPR006029; Neu channel memb.
 DR InterPro; IPR006202; Neur chan IBD.
 DR InterPro; IPR006201; Neur channel.
 DR Pfam; PF02931; Neur_chan_IBD; 1.
 DR Pfam; PF02932; Neur_chan_memb; 1.
 DR PRINTS; PR00252; NRIONCHANNEL.
 DR TIGRFAMs; TIGR00860; LIC; 1.
 DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
 DR Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 KW Transmembrane; Multigene family.
 FT SIGNAL 1 22 BY SIMILARITY.
 FT CHAIN 23 502 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,
 FT ALPHA-7 CHAIN.
 FT DOMAIN 23 230 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 231 255 POTENTIAL.
 FT TRANSMEM 262 280 POTENTIAL.
 FT TRANSMEM 296 317 POTENTIAL.
 FT DOMAIN 318 469 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 470 490 POTENTIAL.
 FT DISULFID 150 164 BY SIMILARITY.
 FT DISULFID 212 213 ASSOCIATED WITH RECEPTOR ACTIVATION
 (BY SIMILARITY).
 FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 502 AA; 56631 MW; C9312E5226D120E3 CRC64;
 Query Match 46.6%; Score 1241; DB 1; Length 502;
 Best Local Similarity 47.1%; Pred. No. 1.4e-96;
 Matches 243; Conservative 90; Mismatches 139; Indels 44; Gaps 6;
 QY 1 MCGRARRSHLAAPAGALLLLCLLWPRGARGCVGHEKLLHLLHHLLHHYVNLPRPVNESDPLQ 60
 Db 1 MCGRRGGIWLALAAALLHV-----SLQGFQRRLYKELVKNYPLERPVANDSQPLT 52
 QY 61 LSPGLTLMQIIVDDEKNOLLITNMLKLEWDMNLRWNTSDFGVKDLRVPPHRLWKPDV 120
 Db 53 VYFSLQLQIMVDDEKNQVLTINMLQMSWTDHYLQNNMSEYGVKNVRFPDQIWKPD 112
 QY 121 LMVNSADEGFDSTPTNVVVRNNGSCLYVPPGIFKSTCKIDITWFPDDQRCMKFGSWT 180
 Db 113 LLYNSADERDFTHTNLVNASGHQCYLPPIGFKSSCYIDVRWFPDVGQCKLKFGSWS 172
 QY 181 YDGYQLDLQDEGGDISFVTVNGEWELIGVGKNEIYNNCPBPYIDITFAVVRK 240
 Db 173 YGGWSLDLQOE---ADISSYIENGWDLMGIFGKNEKEFEYCKEYPDVTYVTVNRR 229
 QY 241 TLYYFENLIVPCVLIASMLLGTLPDSCGKLSGLVTLLSLTVFLNVAETMPATSDA 300
 Db 230 TLYYGLNLLIPCVLISALALLVFLPADSGEKISLGTIVLLSITVFMILVAEIMPASDS 289
 QY 301 VPLLGTVFCNMFVAVSSVSTLIILNHHRRHADTHEMSDWIRCVFLYMLPWLRLMRSPG 360
 Db 290 VPLIAQYFASTMIIVGLSVVTVIVLRYHHHDPDGGKMPKWTIRVILLNCAWFLMRKRP 349
 QY 361 SATTPPARVPPDLELRERSSKSLIANVLDIDDDFRHQAQPPCCRYRGGE----- 415
 Db 350 EDKVRPACQHKP-----RRCSLASV-ELSGAGAPPTSGNLLYIGRGLGEMHCA 398
 QY 416 ---ENGAGLAHSCFGV-----DYLSLILKEIRVITDQMRKDDADISR 458
 Db 399 PTPDGGVVCGRMACSPHDEHLMGTHPSDGDPLAKILBEVRYIANRFRCODESEVICS 458
 QY 459 DWKFAAMVVDRLCLIIFTLTIIATLAVLSAPHIM 494
 Db 459 EWKFAACVVDRLCLMAFSVFTIICTIGILMSAPNFV 494

RESULT 4
 ACH7_RAT STANDARD; PRT; 502 AA.
 ID ACH7_RAT
 AC Q05941;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neuronal acetylcholine receptor protein, alpha-7 chain precursor.
 GN CHRNA7 OR ACRA7.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93147931; PubMed=7678857;
 RA Seguela P., Wadiche J., Dineley-Miller K., Dani J.A., Patrick J.W.;
 RT "Molecular cloning, functional properties, and distribution of rat

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

EMBL; X39364; CAA53802.1; --
DR InterPro; IPR006029; Neur_chan_memb.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006201; Neur_chan.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMS; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane; Multigene family; Alternative splicing.
FT SIGNAL 1 19
FT CHAIN 20 499
FT ALPHA-7 CHAIN.
FT EXTRACELLULAR (POTENTIAL).
FT DOMAIN 20 227
FT TRANSMEM 228 252
FT TRANSMEM 259 277
FT TRANSMEM 293 314
FT DOMAIN 315 466
FT TRANSMEM 467 487
FT DISULFID 147 161
FT DISULFID 209 210
FT BY SIMILARITY.
FT ASSOCIATED WITH RECEPTOR ACTIVATION
FT (BY SIMILARITY).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 43 43
FT CARBOHYD 87 87
FT CARBOHYD 130 130
FT VARSPLIC 262 290
FT Missing (in isoform Short).
FT /FTId=VSP_000075.
FT SEQUENCE 499 AA; 56002 MW; AEE5D0B3820D42D5 CRC64;

Query Match 45.7%; Score 1218.5; DB 1; Length 499;
Best Local Similarity 46.1%; Pred. No. 1e-94;
Matches 241; Conservative 80; Mismatches 115; Indels 87; Gaps 9;

QY 26 RGARC-----GYHEKRLHLLHLDHYNNLPRPVNNSDPLQLSGLTLMQIID 72
DB 2 RGSICLALASILHVSLOGEFQKLYKDLVKNYFLERPVANDSLPTVVFSLSLQIMD 61
QY 73 VDEKNQLITNIMKLEWDMNLRWNTSDFGGVKDLRVPHRLWKPDVLMVNSADEGDS 132
DB 62 VDEKNQVLTNIMQMTWTDHYLQWNASPYGVKTVRFPDQGIWKPDIILYNSADERFDA 121
QY 133 TYPTNVVRNGSLVYPPGIFKSTCKIDITWRPFDDQRCMEKFGSWTYDGYQLDLQLOD 192
DB 122 TFTNVLVNSGHCQYLPGLFSSCYIDVRWPFDDVQCKLFGSGWSYGGWLDLQOE 181
QY 193 EGGGDISFVTNGEWELIGVPGKNEYIYNCCEPPYIDITPAVVRKTLTYFFNLVIPC 252
DB 182 ---ADISGYIPNGEWDVGLVKGSEKFECCKEPPYDVTFTVSIRRTLYYGLNLLIPC 238
QY 253 VLTASMLLGTLPDPSGKSLGVTLLSLTFLNVAETMPATSDAVPLLGTYFNCIM 312
DB 239 VLISALALLVFLPADSGEKISLGITVLLSLTFLVEMLLVAETMPATSDVPLIAYFASTM 298
QY 313 FMVASSVSTLLIYNHHRHADTHEMSDWTRCVFLYWLPLVLRMSRPD----- 360
DB 299 IIVGLSVVTVIVIOYHHHDPDGKMPKNTVRVLLNWCAMFLMKRGEDKVRACQHNE 358
QY 361 -----SATTPPPA-----RVPPPPPLELRSSKSLLANVL--- 391
DB 359 RRCSLASVMSAVGPPATNGNLLYIGRGLDTHMCAPTDP-----SGVCCGRVACSP 411
QY 392 DDDDFRHPQAQQCCCPATNGYRGGEENGAGLAHSCFGVDYVLSLLKLEIRVITQMKRDD 451

DB 412 TDEHLLH--AQGP-----SEG-----PDLAKILEEVRYIAHRFCQD 448

QY 452 EDADISRDWKFAMVVDRLCLIFTLFTIATLAVLLSAPHIM 494
DB 449 ESEAVCSEWKFACVVDRLCLMAFVFTILCTIGILMSAPNFV 491

RESULT 6
ACH1 CAEL
ID ACH1 CAEL STANDARD; PRT; 498 AA.
AC P48180;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acetylcholine receptor like protein, alpha-type chain precursor.
GN P25G6.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=96196478; PubMed=8627624;
RA Ballivet M., Alliod C., Bertrand S., Bertrand D.;
RT "Nicotinic acetylcholine receptors in the nematode Caenorhabditis
RT elegans.";
RL J. Mol. Biol. 258:261-269(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=96196478; PubMed=8627624;
RA Nelson J., Wohldmann P.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Possible acetylcholine receptor.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the ligand-gated ionic channel family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

EMBL; X83887; CAA58764.1; --
DR EMBL; AF022973; AAC25796.1; --
DR PIR; S68588; S68588.
DR HSP; P58154; I19B.
DR Wormpep; F25G6.3; CE09639.
DR InterPro; IPR006029; Neur_chan_memb.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006201; Neur_chan.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMS; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein;
KW Transmembrane; Signal.
FT SIGNAL 1 19
FT CHAIN 20 498
FT POTENTIAL.
FT ACETYLCHOLINE RECEPTOR LIKE PROTEIN,
FT ALPHA-TYPE CHAIN.
FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 231 252
FT TRANSMEM 261 279
FT TRANSMEM 295 314
FT TRANSMEM 315 472
FT DOMAIN 473 493
FT TRANSMEM 473 493
FT DISULFID 147 161
FT DISULFID 211 212
FT CARBOHYD 43 43

RA Elliott K.J., Ellis S.B., Berckhan K.J., Urrutia A.,
 RA Chavez-Noriega L.E., Johnson E.C., Velicelebi G., Harpold M.M.;
 RT "Comparative structure of human neuronal alpha 2-alpha 7 and beta
 RT 2-beta 4 nicotinic acetylcholine receptor subunits and functional
 RT expression of the alpha 2, alpha 3, alpha 4, alpha 7, beta 2, and
 RT beta 4 subunits.";
 RL J. Mol. Neurosci. 7:217-228(1996).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=97162233; PubMed=9009220;
 RA Groot Kormelink P.J., Luyten W.H.M.L.;
 RT "Cloning and sequence of full-length cDNAs encoding the human neuronal
 RT nicotinic acetylcholine receptor (nAChR) subunits beta3 and beta4 and
 RT expression of seven nAChR subunits in the human neuroblastoma cell
 RT line SH-SY5Y and/or IMR-32";
 RL FEBS Lett. 400:309-314(1997).
 RN [5]
 RP SEQUENCE FROM N.A., AND VARIANT LEU-21 INS.
 RX MEDLINE=21342809; PubMed=11450844;
 RA Lev-Lehman E., Bercovich D., Xu W., Stockton D.W., Beaudet A.L.;
 RT "Characterization of the human beta4 nAChR gene and polymorphisms in
 RT CHRNA3 and CHRNA4";
 RL J. Hum. Genet. 46:362-366(2001).
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo T.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uddin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fabley J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [8]
 RP SEQUENCE OF 30-503 FROM N.A.
 RX TISSUE=Brain;
 RA Anand R., Lindstrom J.;
 RL Submitted (JUN-1990) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP SEQUENCE OF 6-493 FROM N.A.
 RX TISSUE=Epidermal keratinocytes;
 RA Arredondo J., Grando S.A.;
 RT "Cloning cholinergic receptors in human keratinocytes.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
 CC extensive change in conformation that affects all subunits and
 CC leads to opening of an ion-conducting channel across the plasma
 CC membrane.
 CC -!- SUBUNIT: Neuronal AChR seems to be composed of two different type
 CC of subunits: alpha and non-alpha (beta).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=P32297-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=P32297-2; Sequence=VSP_000073;
 CC Note=No experimental confirmation available;
 CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M864176.1; AAC84176.1; -;
 CC EMBL; M37981; AA559942.1; -;
 CC EMBL; U62432; AAB40110.1; -;
 CC EMBL; Y084418; CAA69695.1; -;
 CC EMBL; AJ007783; CAA07682.1; -;
 CC EMBL; AJ007784; CAA07682.1; JOINED.
 CC EMBL; AJ007785; CAA07682.1; JOINED.
 CC EMBL; AJ007786; CAA07682.1; JOINED.
 CC EMBL; AJ007787; CAA07682.1; JOINED.
 CC EMBL; BC001642; AAH01642.1; -;
 CC EMBL; BC002996; AAH02996.1; -;
 CC EMBL; BC000513; AAH00513.1; -;
 CC EMBL; AF385584; AAK68110.1; -;
 CC EMBL; X53559; CAA37625.1; -;
 CC PIR; A37040; A37040.
 CC PIR; A53956; A53956.
 CC Genew; HGNC:1957; CHRNA3.
 CC MIM; 118503; -;
 CC GO; GO:0005892; C:nicotinic acetylcholine-gated receptor-chan. . . ; TAS.
 CC GO; GO:0015464; F:acetylcholine receptor activity; TAS.
 CC GO; GO:0004889; F:nicotinic acetylcholine-activated cation-se. . . ; TAS.
 CC GO; GO:0005215; F:transport activity; TAS.
 CC GO; GO:0007165; F:signal transduction; TAS.
 CC GO; GO:0006810; P:transport; TAS.
 CC InterPro; IPR006029; Neu_chan_LBD.
 CC InterPro; IPR006202; Neu_chan_LBD.
 CC InterPro; IPR006201; Neu_chan_LBD.
 CC Pfam; PF02931; Neur_chan_LBD; 1.
 CC PRINTS; PR00252; NRIONCHANNEL.
 CC TIGRFAMs; TIGR00860; LIC; 1.
 CC PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
 CC Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 CC Transmembrane; Multigene family; Alternative splicing; Polymorphism.
 CC SIGNAL 1 29
 CC CHAIN 30 503
 CC
 CC DOMAIN 30 238
 CC TRANSMEM 239 263
 CC TRANSMEM 271 289
 CC TRANSMEM 305 326
 CC DOMAIN 327 475
 CC TRANSMEM 476 495
 CC DISULFID 157 171
 CC DISULFID 221 222
 CC
 CC CARBOHYD 53
 CC CARBOHYD 170 170
 CC VARSPLIC 1 5
 CC
 CC VARIANT 21 21
 CC
 CC CONFLICT 5 13
 CC CONFLICT 11 14
 CC CONFLICT 100 100
 CC CONFLICT 132 133
 CC CONFLICT 235 235

Db 149 NNADGEFAVTHMTKAHLFGTGTGVHVPVPAIYKSCSIDVTFFPDQNCQKMGFSWTYDK 208
 QY 184 YQLDLQLOBGGDISFVINGEWELLIGVQKNEIYVNCPPYDITFAVVRKTLTY 243
 Db 209 AKIDLE-QMEQTVLDKYWESGWAIVNAGTGNKSKYDCCAIYPDVYAFVIRLPLF 267
 QY 244 YFFNLIVPCVLIIASMAALLGFTLPDGEKLSGLGVTLLSLTVFLNMVAETMPATSDAVPL 303
 Db 268 YTINLIIPCLLISCLTVLVEYFSDGCKTLCISVLLSTVFLLEITLIPSTSLVPL 327
 QY 304 LGFYNCIMPVASSVSTILINYYHHRHADTHEMSDWIRCVLYLWPLWVLRMSRGSAT 363
 Db 328 IGEYLLFTMTFVTLISVITVFLNVHRSPTSTMPHWRGALLGCVPRLLMNRP---- 383
 QY 364 TPPARVPPPPDLELRSSKSLIANVLD-----IDDDFR----- 398
 Db 384 -PPFVELCHP--LRLKSLPSYHLENSVDAEREVVVEEDRWACAGHVAPSVTILCSHG 440
 QY 399 --HPOAQQPCQCRYYRGEENGAGLAHSCFGVDYELSLILKEIRVITDQMRKDDDEDADI 456
 Db 441 HLHSGASGPKAEALLQEGE--LLLSPH-----MQKALEGVHYIADHLRSEDADSSV 489
 QY 457 SRDWKFAAMVVRCLLIITFLTIATLAVLL 488
 Db 490 KEDWKYVAMVIDRIFLWLFIVCFGLGTIGLFL 521

RESULT 11
 ACH2 CHICK
 ID_ACH2_CHICK STANDARD; PRT; 528 AA.
 AC P09480;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neuronal acetylcholine receptor protein, alpha-2 chain precursor.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=88283624; PubMed=3267226;
 RA Nef P., Onysier C., Alliod C., Coururier S., Ballivet M.;
 RT "Genes expressed in the brain define three distinct neuronal
 RT nicotinic acetylcholine receptors.";
 RL EMBO J. 7:595-601(1988).
 CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
 CC extensive change in conformation that affects all subunits and
 CC leads to opening of an ion-conducting channel across the plasma
 CC membrane.
 CC -!- SUBUNIT: Neuronal AChR seems to be composed of two different
 CC type of subunits: alpha and non-alpha (also called beta). A
 CC functional receptor seems to consist of two alpha-chains and
 CC three non-alpha chains.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; X07339; CAB59645.1; -
 CC EMBL; X07340; CAB59645.1; JOINED.
 CC EMBL; X07341; CAB59645.1; JOINED.
 CC EMBL; X07342; CAB59645.1; JOINED.
 CC EMBL; X07343; CAB59645.1; JOINED.
 CC EMBL; X07344; CAB59645.1; JOINED.

DR EMBL; AJ250360; CAB59625.1; -.
 DR PIR; S00377; ACCH2N.
 DR InterPro; IPR006029; Neu_channel_memb.
 DR InterPro; IPR006202; Neur_chan_IIBD.
 DR InterPro; IPR006201; Neur_channel.
 DR Pfam; PF02931; Neur_chan_IIBD; 1.
 DR Pfam; PF02932; Neur_chan_memb; 1.
 DR PRINTS; PR00252; NRIONCHANNEL.
 DR TIGRFAMS; TIGR00860; LIC; 1.
 DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
 KW Postsynaptic membrane; Neuronal ION CHANNEL; Signal;
 KW Transmembrane; Multigene family.
 FT SIGNAL 1 23
 FT CHAIN 24 528 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,
 FT ALPHA-2 CHAIN.
 FT EXTRACELLULAR.
 FT DOMAIN 24 239
 FT TRANSMEM 240 264
 FT TRANSMEM 272 290
 FT TRANSMEM 306 327
 FT DOMAIN 328 501
 FT TRANSMEM 502 520
 FT DOMAIN 394 410
 FT DISULFID 158 172
 FT DISULFID 222 223
 FT CARBOHYD 54 54
 FT CARBOHYD 104 104 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT SEQUENCE 528 AA; 60675 MW; E76C6360AF876364 CRC64;
 SQ
 Query Match 36.0%; Score 959; DB 1; Length 528;
 Best Local Similarity 39.2%; Pred. No. 7.2e-73;
 Matches 198; Conservative 84; Mismatches 161; Indels 62; Gaps 10;
 QY 31 GXHEKRLHLHLHDHNVNLERPVVNESDPLQLSFGLTLMQIIDVDEKNQLITINIKLEW 90
 Db 31 GFAEDRLFKHFTGYNRWSRPVNTSDVIVVFGLSIAQLIDVDEKNQMTINVLKQEW 90
 QY 91 NDMNLEWNTSDGGYKDLVPPHRLWKPDVLMYNSADEGFDSTYPTNVVVRNNGSLYVP 150
 Db 91 SDYKLRNPEDFDNVTISRVPSEMIIPDIVLYNNADGEFAVTHMTKAHLFSGNKVKKVP 150
 QY 151 PGIFKSTCKIDITWRPFDQRCMKFGSWTYDGYQLDLQLOBGGDISFVINGEWELI 210
 Db 151 PAIYKSSCSIDVTYFFDQNCMKFGSWTYDKAKIDLE-NMEHHVDLKDYESGEWALI 209
 QY 211 GVPGRKEIYVNCPEPYDITFAVVRKTLTYFFNLIVPCVLIIASMAALLGFTLPDSDG 270
 Db 210 NAIGRYSKKYDCCTEYDITFYFVIRRLPLFTYINLIIPCLLISCLTVLVEYFPLSDCG 269
 QY 271 EKLGLVTLISLTVFLNMVAETMPATSDAVLLGTYFNCIMPVASSVSTILINYYH 330
 Db 270 EKITLCISVLLSLTVFLLEITLIPSTSLVLEIGEYLLFTMTFVTLISITVFLNVH 329
 QY 331 READTHEMSDWIRCVLYLWPLWVLRMSRGSATTPPARVP-----PPDLELRERS 382
 Db 330 RSPSTHWPWVRSFPLGFIPLWPKR-----PPLLLPAEGTQYDPPGTRL--ST 380
 QY 383 SKSLIANVLDIDDDFRHPQAOQPQ-----CCRYRGEENGAGLAH-SC----- 426
 Db 381 SRCWLET--DVDDKWESEEEEEEEEEEEBKAYPKRVPSGSGQ--GTOCHYSCERQA 435
 QY 427 -----FGVDVLSI-----ILKEIRVITDQMRKDDDEDADI SRDWKFA 463
 Db 436 GKASGGAPQVPLKGEVSGSQGLTSPSILRALEGVQVIADHLRAEDADFSVKEDWKYV 495
 QY 464 AMVDRCLLIITFLTIATLAVLL 488
 Db 496 AMVIDRIFLWFLIIVCLLGTGVLFL 520
 RESULT 12
 ACH1_DROME
 ID_ACH1_DROME STANDARD; PRT; 567 AA.


```

Db 249 IPCWGISFLSVLYFLPSDGEKISLCISILLSTLVFFLLAELIIPPTSLTVLLGKYL 308
QY 310 CIMFVASSVSTILLINHYHRHADTHEMDWIRCVFLYWLFWLWLRMSRPGSATPPPAR 369
Db 309 FTMLVLSVVVTVIIVNLNFRSPVTHRMAPVQRLFIQLPKLCTIERP--KKEEPEED 366
QY 370 VPP-----PPDLERLRSKSL-----L 387
Db 367 QPPEVLTVVHLPPDVVDKFNVDKSPGSDYGIPLPASHRFDLAAGGISAHCFAEPL 426
QY 388 ANVLDI---DDDFRHPQAO---QPOCCRYRGGEENGAGLAHRS---CFGVYD--ELS 434
Db 427 PSSPLPGADDDLFSPSLGNDISPGCC-----PAAAAAADLSPTFEXPYAREME 478
QY 435 LILKEIRVITDQMRKDEDEDADISRDNKFAMVVDRLCLIFTFTIATLAVLSAPHI 493
Db 479 KTEGSRFIAHQVKNKDFESVEDWKYVAVLDRMFLMFAICAVVGTALIILOAPSL 537

RESULT 13
ACH2_RAT
ID ACH2_RAT STANDARD; PRT; 511 AA.
AC F12389; 008952;
DT 01-OCT-1989 (Rel. 12, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-2 chain precursor.
GN CHRNA2 OR ACRA2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=88178113; PubMed=2832952;
RA Wada K., Ballivet M., Boulter J., Connolly J.G., Wada E.,
RA Deneris E.S., Swanson L.W., Heinemann S.F., Patrick J.;
RT "Functional expression of a new pharmacological subtype of brain
RT nicotinic acetylcholine receptor."
RL Science 240:330-334 (1988).
RN [2]
RP REVISIONS.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RA Boulter J.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
CC extensive change in conformation that affects all subunits and
CC leads to opening of an ion-conducting channel across the plasma
CC membrane.
CC -!- SUBUNIT: Neuronal AChR seems to be composed of two different types
CC of subunits: alpha and non-alpha (beta). Alpha-2 subunit can be
CC combined to beta-2 or beta-4 to give rise to functional receptors.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; L10077; AAB60900.1; -
CC EMBL; M20297; AAA40664.1; -
CC EMBL; M20292; AAA40664.1; JOINED.
CC EMBL; M20293; AAA40664.1; JOINED.
CC EMBL; M20294; AAA40664.1; JOINED.
CC EMBL; M20295; AAA40664.1; JOINED.
CC EMBL; M20296; AAA40664.1; JOINED.
CC EMBL; M20297; AAA40664.1; JOINED.
CC InterPro; IPR006029; Neu_channel_memb.

```

```

DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006201; Neur_channel.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMS; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane; Multigene family.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 511 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,
FT DOMAIN 28 241 ALPHA-2 CHAIN.
FT TRANSMEM 242 266 EXTRACELLULAR.
FT TRANSMEM 274 292 POTENTIAL.
FT TRANSMEM 308 329 POTENTIAL.
FT TRANSMEM 330 484 CYTOPLASMIC.
FT TRANSMEM 485 503 POTENTIAL.
FT DOMAIN 391 402 POLY-GLU.
FT DISULFID 160 174 BY SIMILARITY.
FT DISULFID 224 225 ASSOCIATED WITH RECEPTOR ACTIVATION
FT CARBOHYD 56 56 (BY SIMILARITY).
FT CARBOHYD 106 106 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 494 494 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 494 494 C -> S (IN REF. 1; AAA40664).
SQ SEQUENCE 511 AA; 58611 MW; 3824E83BB01D613B CRC64;

Query Match 35.7%; Score 951; DB 1; Length 511;
Best Local Similarity 39.4%; Pred No. 3.3e-72;
Matches 198; Conservative 90; Mismatches 167; Indels 48; Gaps 8;

QY 16 LLLLLCLMPRGARCGYHEKRLHLHLHDHYNLVRPVVNESDPLQLSFLGLTLMQIIDVE 75
Db 19 LLLVPAVLTVQGSHT-HAEDRLFKLFGCYNWARVPNTSDVIVRFGLSIAQLIDVE 77
QY 76 KNOILLTNLWKLWNMDNLRNNTSDFGKXKDLRVPPLWKPDLVLMYNSADEGFSTYP 135
Db 78 KQMMTTNVLKQEWNDYKLRWDPAEFGNVTSLRVPSEMIWIPDIVLYNNADGEFAVTM 137
QY 136 TNVWVRNNGSLYVPPGIFKSTCKIDITWFPDDQRCMKFGSWTVGYQLDLQLQDEGG 195
Db 138 TXAHLLFTGTWVPPAIYKSSCSIDVFFPDQNCMKFGSWTVYDKAKIDLE-QMERT 196
QY 196 GDISSFVTNGEWELIGVPGKRNEIYNCPEPIDITFAVIRRKTYLYFFNLIVPCVLI 255
Db 197 VDLKDYWESGEWAIINATGTYSKKYDCCAEIYPDVYVYFVIRRLPFLFTINLIIFCLLI 256
QY 256 ASWALLGFTLPPDSGKLSGLVTILLSLTVLMVAETMPATSDAVPLGLTYFNCIMFWY 315
Db 257 SCLTVLVFLPSECGEKITLCISVLSLTVFLLITEIIPSTSLVIPLIGEYLLFTMIV 316
QY 316 ASSVSTIILINHYHRHADTHEMSDWIRCVFLYWLFWLWLRMSRPGSATPPPARVPPPD 375
Db 317 TLSIVITVFLNVHRSSTHNPWVRVALLGRVPRWLMNRP-----LPPWELHGSPP 371
QY 376 LEIERSKSLANVLDDDFRHPQAOQPOCCRYRGGEEN---GAGLAHSCFGVDY- 431
Db 372 LKL--SPSVHWTETNMDAGERETEETEEEE-----DENICVAGL-PDSSMGVLVG 420
QY 432 -----ELSLILKEIRVITDQMRKDEDEDADISRDNKFAMVVDRLCLIFTFTIATLAVLSAPHI 465
Db 421 HGGLHLRAMEPTKTPSQASEILLSPQIKALEGVHYIADRLRSEDADSSVKEDWKYVAM 480
QY 466 VVDRCLLIIFTFTIATLAVLL 488
Db 481 VVDRFLWLFIVFCFLGIGLFL 503

RESULT 14
ACH3_BOVIN
ID ACH3_BOVIN STANDARD; PRT; 495 AA.
AC Q07263;

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:30:03 ; Search time 31.1579 Seconds

(without alignments)
5022.709 Million cell updates/sec

Title: US-09-303-232-4

Perfect score: 2665

Sequence: 1 MCGRRRSHLAAPAGLLLLL.....LFTIATLAVLLSAPHIMWS 496

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2665	100.0	496	5 Q9XZI3	Q9xzi3 heliothis v
2	2064.5	77.5	807	5 Q8T7V5	Q8t7v5 drosophila
3	2047	76.8	542	5 Q8GMN7	Q8gm7 drosophila
4	2046.5	76.8	545	5 Q9VWI9	Q9vwi9 drosophila
5	1803.5	67.7	501	5 Q9XZI4	Q9xzi4 heliothis v
6	1795.5	67.4	494	5 Q8T7S2	Q8t7s2 drosophila
7	1792.5	67.3	494	5 Q8T7S1	Q8t7s1 drosophila
8	1791.5	67.2	494	5 Q8GMN8	Q8gm8 drosophila
9	1790.5	67.2	494	5 Q8T7S3	Q8t7s3 drosophila
10	1774	66.6	509	5 Q8T7S0	Q8t7s0 drosophila
11	1769	66.4	523	5 Q8T7R9	Q8t7r9 drosophila
12	1484.5	55.7	554	5 Q9VLT9	Q9vlt9 drosophila
13	1371	51.4	525	5 Q8LPE2	Q8lpe2 drosophila
14	1272.5	47.7	509	13 Q8Q0C7	Q8q0c7 brachydanio
15	1253	47.0	555	13 Q7T2U0	Q7t2u0 fugu rubrip
16	1251	46.9	570	5 Q9VJT9	Q9vjt9 drosophila

17	1242	46.6	502	11	Q9JHD6	Q9jhd6 mus musculu
18	1239.5	46.5	486	13	Q7T2S0	Q7t2s0 fugu rubrip
19	1235	46.3	502	6	Q866A2	Q866a2 macaca mula
20	1219.5	45.8	511	13	Q03481	Q03481 gallus gall
21	1207	45.3	513	13	Q7T2R9	Q7t2r9 fugu rubrip
22	1166.5	43.8	554	13	Q7T2T9	Q7t2t9 fugu rubrip
23	1150.5	43.2	474	13	Q7T2U1	Q7t2u1 fugu rubrip
24	1141	42.8	461	5	P91197	P91197 caenorhabdi
25	1113	41.8	480	5	Q8I932	Q8i932 caenorhabdi
26	1106.5	41.5	335	5	Q9NKD1	Q9nkd1 drosophila
27	1058.5	39.7	554	5	O62083	O62083 caenorhabdi
28	1055.5	39.6	542	5	O18556	O18556 caenorhabdi
29	994	37.3	523	5	O46128	O46128 heliothis v
30	975.5	36.6	537	5	Q9U941	Q9u941 myzus persi
31	972	36.5	515	5	O46133	O46133 locusta mig
32	962	36.1	537	5	Q8MUR0	Q8mur0 apis mellif
33	961	36.1	505	4	Q8U077	Q8u077 homo sapien
34	960	36.0	536	5	Q8T0Y9	Q8t0y9 aplysia cal
35	956.5	35.9	509	5	Q8MM21	Q8mm21 aphid gossy
36	951	35.7	536	5	Q8T9S0	Q8t9s0 aplysia cal
37	949.5	35.6	509	5	Q9NFX8	Q9nfx8 myzus persi
38	947	35.5	568	5	Q9NFR5	Q9nfr5 drosophila
39	942.5	35.4	552	5	P91765	P91765 myzus persi
40	940.5	35.3	512	11	O91X60	O91x60 mus musculu
41	937.5	35.2	545	5	O96631	O96631 heliothis v
42	932	35.0	499	11	Q8VHH6	Q8vhh6 mus musculu
43	929	34.9	540	5	O46134	O46134 locusta mig
44	928	34.8	499	11	Q8R4G9	Q8r4g9 mus musculu
45	928	34.8	504	11	Q8BV44	Q8bv44 mus musculu

ALIGNMENTS

RESULT 1

Q9XZI3

ID Q9XZI3 PRELIMINARY; PRT; 496 AA.

AC Q9XZI3; (TREMREL. 12, Created)

DT 01-NOV-1999 (TREMREL. 12, Last sequence update)

DT 01-JUN-2003 (TREMREL. 24, Last annotation update)

DE Putative nicotinic acetylcholine receptor alpha 7-1 subunit.

OS Heliothis virescens (Noctuid moth) (Owlet moth).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Noctuoidea; Noctuidae; Heliothinae; Heliothis.

OX NCBI_TaxID=7102;

RN [1]

RP SEQUENCE FROM N.A.

RA Schulte T., Oellers N., Adamczewski M.;

RT "Putative alpha subunits of insect nicotinic acetylcholine receptors more similar to vertebrate alpha 7 subunits and C. elegans Ce21 than to other insect nicotinic acetylcholine receptor alpha subunits.";

RL Submitted (APR-1999) to the EMBL/GenBank/DDBI databases.

CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.

DR EMBL; AF143846; AAD32697.1; -.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti...; IEA.

DR GO; GO:0005216; F:ion channel activity; IEA.

DR GO; GO:00030594; F:neurotransmitter receptor activity; IEA.

DR GO; GO:0006811; P:ion transport; IEA.

DR GO; GO:0007269; P:synaptic transmission; IEA.

DR InterPro; IPR006201; Neur channel.

DR InterPro; IPR006202; Neur_chan_LBD.

DR InterPro; IPR006029; Neur_chan_LBD_memb.

DR Pfam; PF02931; Neur_chan_LBD; 1.

DR Pfam; PF02932; Neur_chan_memb; 1.

DR PRINTS; PRO0252; NRIONCHANNEL.

DR TIGRFAMs; TIGR00860; LIC; 1.

DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.

KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor; Transmembrane.

```

SQ SEQUENCE 496 AA; 56347 MW; 8032FED8515A6210 CRC64;
Query Match 100.0%; Score 2665; DB 5; Length 496;
Best Local Similarity 100.0%; Pred. No. 9.4e-251;
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGRARRSHLAAPAGALLLCLLPARGCGYHEKRLHLLDHYNVLPRPVVNESDPLQ 60
DB 1 MGRARRSHLAAPAGALLLCLLPARGCGYHEKRLHLLDHYNVLPRPVVNESDPLQ 60
QY 61 LSFGLTLMQIIVDEKNQLLITNIWLEWNDMNLWNTSDFGVKDLRVPVHRLWKPDV 120
DB 61 LSFGLTLMQIIVDEKNQLLITNIWLEWNDMNLWNTSDFGVKDLRVPVHRLWKPDV 120
QY 121 LMTNSADEGSDSYPTNVTNRNGSCLYVPPGIFKSTCKIDITWFPDDQRCCKFGSWT 180
DB 121 LMTNSADEGSDSYPTNVTNRNGSCLYVPPGIFKSTCKIDITWFPDDQRCCKFGSWT 180
QY 181 YDGVQLDLOQDEGGDISFVTNGEWELIGVPGKNEIYVNCPEYIDITFAVVIIRK 240
DB 181 YDGVQLDLOQDEGGDISFVTNGEWELIGVPGKNEIYVNCPEYIDITFAVVIIRK 240
QY 241 TLYYFNLIVPCVLIIASMLLGTLPDPSGKLSLGVITLLSLTVFLNMVAETMPATSDA 300
DB 241 TLYYFNLIVPCVLIIASMLLGTLPDPSGKLSLGVITLLSLTVFLNMVAETMPATSDA 300
QY 301 VPLLGTFCNIMPVASSVSTILINHYHRRHADTHEMSDWIRCVFLYMLPWVLRMRPG 360
DB 301 VPLLGTFCNIMPVASSVSTILINHYHRRHADTHEMSDWIRCVFLYMLPWVLRMRPG 360
QY 361 SATTPPPARYPPPPDLRLRRSSKSLANVLIDDDFRHPQAOQPCCRYYRGGEEAG 420
DB 361 SATTPPPARYPPPPDLRLRRSSKSLANVLIDDDFRHPQAOQPCCRYYRGGEEAG 420
QY 421 LAHSCFGVYELSLILKEIRVITDQMKDEDAISRDWKFAAMVVDRLCLIFTFTI 480
DB 421 LAHSCFGVYELSLILKEIRVITDQMKDEDAISRDWKFAAMVVDRLCLIFTFTI 480
QY 481 IATLAVLLSAPHIMVS 496
DB 481 IATLAVLLSAPHIMVS 496

RESULT 2
Q87TV5 PRELIMINARY; PRT; 807 AA.
ID Q87TV5 AC Q87TV5
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nicotinic acetylcholine receptor Dalphas5 subunit.
GN NACR-ALPHA-34E OR NACRALPHA-34E OR BG:DS05899.4 OR BG:DS05899.5 OR
CG4498 OR CG16878 OR CG32975.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalphas5, Dalphas6 and Dalphas7, in Drosophila melanogaster Identify a
RT New and Highly Conserved target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.";
RT Genetics 160:1519-1533(2002).
RC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF272778; AAM13390.1; -.
DR FlyBase; FBgn028875; nAcR-alpha-34E.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
```



```

RESULT 6
Q8T7S2
ID Q8T7S2 PRELIMINARY; PRT; 494 AA.
AC Q8T7S2;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Nicotinic acetylcholine receptor Dalphaf subunit variant type II.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalphaf, Dalphaf and Dalphaf7, in Drosophila melanogaster Identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing."
RL Genetics 160:1519-1533 (2002).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321446; AAM13393.1; -
DR FlyBase; FBgn0032151; nacr-alpha-30D.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005230; F: extracellular ligand-gated ion channel acti. . .; IEA.
DR GO; GO:0005216; F: ion channel activity; IEA.
DR GO; GO:0030594; F: neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P: ion transport; IEA.
DR GO; GO:0007268; P: synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neu_chan_memb.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRfam; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 494 AA; 56048 MW; 6EE711810E0DE7BBB CRC64;

Query Match 67.4%; Score 1795.5; DB 5; Length 494;
Best Local Similarity 68.1%; Pred. No. 5.3e-166;
Matches 342; Conservative 54; Mismatches 81; Indels 25; Gaps 5;

QY 8 SHLAAPAGLLLLCLLWPRGARC-GVHEKRLHLLDHNVLPRVPPVNESDPLQLSFGLT 66
Db 3 SLPASLSLFLVLLIFAIKESCGQPHKRLNLLSTNTLPRVANESEPLEVFGFLT 62
QY 67 LMQIIDVDEKQQLITNIWLKLEWMDNMLRWNTSDFGVKDLRVPVPHRLWKPDVLMYNSA 126
Db 63 LQOIIDVDEKQQLITNLWLSLEWMDNMLRWNETYGVKDLRITNKLWKPDVLMYNSA 122
QY 127 DEGFSTYTNVVRNGSLYVPPGIFKSTCKIDITWFFDDQRCMKFGSWTYDGYQL 186
Db 123 DEGFDTGTYTNIVVRHGSCLYVPPGIFKSTCKMDITWFFDDQRCMKFGSWTYDGNQL 182
QY 187 DLQIQDEGGDSSFTVNGEWELIGVPGKRNELTYNCCPEYDITFAVVIRKTLTYFF 246
Db 183 DIVANSEDGGLSDFTINGEWLLAMPKKNTIVVYACCPPEYVDITFTQIRRTLYYFF 242
QY 247 NLIVPCVLIASMALLGFTLPDPSGKLSLGVITLISLTVFLNVAETMPATSDAVELICT 306
Db 243 NLIVPCVLISSMALLGFTLPDPSGKLTGLVITLISLTVFLNVAESMTTSDAVPLICT 302
QY 307 YFNCIMFVASSVSTILINHHHADTHEMSDWIRCVFLYWLWPLVLRMSRPGSATTP 366
Db 303 YFNCIMFVASSVLTIVVYVNLHHRTADTHEMPPWIKSVFLQWLPWLRMGPRGKIKR 362

```

```

Db 123 DEFGDTYHTNVVKGSGCLYVPGIFKSTCKWDITWFPDDOHCMEKFGSWTYDGNQL 182
QY 187 DLOQDEGGDISFVTNGEWELIGVPGKNEIYVNCPPPYDITFVAVIRKTLYYFF 246
Db 183 DLVNGEDGDLSDFTNGEWYLLAMPKKNITVYACCPPEYVDITFTIQRRTLYYFF 242
QY 247 NLIVPCVLIASMLLGTLPDGSSEKLSLGVTLISLTVFLNVAETMPATSDAVPLLGT 306
Db 243 NLIVPCVLISSMALLGTLPDGSSEKLSLGVTLISLTVFLNVAETMPATSDAVPLLGT 302
QY 307 YFCINPMWASSVSTILLINYHRRHADTHEMSDWIRCVFLYWPVLRMSRPGSATPTP 366
Db 303 YFCINPMWASSVSTILLINYHRRHADTHEMSDWIRCVFLYWPVLRMSRPGSATPTP 362
QY 367 PARVP-PPDLELRERSSKSLANVLDDDDFRHPQAOQCCRYRGGEENGAGLAA-- 423
Db 363 TILLSNRMKELEKERSKSLANVLDDDDFRHTIS-----GSQTAIGSSASF 411
QY 424 -----HSCFGVDY-ELSLILKEIRVITDQMRKDEDEDADISRDKWFAAMVVDRLCLI 473
Db 412 GRPTTVEEHTAIGCNHKLHLILKELQFITARMKADDEAELIGDKWFAAMVVDRLCLI 471
QY 474 IFTLFTIIATLAVLLSAPHIMV 495
Db 472 VFTLFTIIATVLLSAPHIIV 493

```

```

RESULT 8
Q86MN8
ID Q86MN8 PRELIMINARY; PRT; 494 AA.
AC Q86MN8
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nicotinic acetylcholine receptor subunit Dalpha6 precursor.
GN NACRALPHA-30D.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Millar N.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ554209; CAD8635.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR InterPro; IPR006201; Neur_chan.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neur_chan_memb.
DR Pfam; PF02932; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW SIGNAL.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 494 NICOTINIC ACETYLCHOLINE RECEPTOR SUBUNIT
FT CHAIN 27 494 DALPHA6.
SQ SEQUENCE 494 AA; 55980 MW; 4DFC572139587070 CRC64;
Query Match 67.2%; Score 1791.5; DB 5; Length 494;
Best Local Similarity 68.1%; Pred. No. 1.3e-165;
Matches 342; Conservative 54; Mismatches 81; Indels 25; Gaps 5;

```

```

QY 8 SHLAAPAGLLLCLLWPRGARC-CYHEKRLIHLHLLHYNVLRPVNPNESDPLQSFGLT 66
Db 3 SPLPASLSFLVLLIFLAIKESCGPHEKRLIHLHLLSYNTLTPRVAESEPVEKFGIT 62

```

```

QY 67 LMQIIDVDEKNQLLITNIWLKLEWNDNLRWNTSDFGGVKDLRVPPHRLWKPDVLMNSA 126
Db 63 LQIIDVDEKNQLLITNIWLKLEWNDNLRWNTSDFGGVKDLRVPPHRLWKPDVLMNSA 122
QY 127 DEFGDTYHTNVVKGSGCLYVPGIFKSTCKWDITWFPDDOHCMEKFGSWTYDGNQL 186
Db 123 DEFGDTYHTNVVKGSGCLYVPGIFKSTCKWDITWFPDDOHCMEKFGSWTYDGNQL 182
QY 187 DLOQDEGGDISFVTNGEWELIGVPGKNEIYVNCPPPYDITFVAVIRKTLYYFF 246
Db 183 DLVNGEDGDLSDFTNGEWYLLAMPKKNITVYACCPPEYVDITFTIQRRTLYYFF 242
QY 247 NLIVPCVLIASMLLGTLPDGSSEKLSLGVTLISLTVFLNVAETMPATSDAVPLLGT 306
Db 243 NLIVPCVLISSMALLGTLPDGSSEKLSLGVTLISLTVFLNVAETMPATSDAVPLLGT 302
QY 307 YFCINPMWASSVSTILLINYHRRHADTHEMSDWIRCVFLYWPVLRMSRPGSATPTP 366
Db 303 YFCINPMWASSVSTILLINYHRRHADTHEMSDWIRCVFLYWPVLRMSRPGSATPTP 362
QY 367 PARVP-PPDLELRERSSKSLANVLDDDDFRHPQAOQCCRYRGGEENGAGLAA-- 423
Db 363 TILLSNRMKELEKERSKSLANVLDDDDFRHTIS-----GSQTAIGSSASF 411
QY 424 -----HSCFGVDY-ELSLILKEIRVITDQMRKDEDEDADISRDKWFAAMVVDRLCLI 473
Db 412 GRPTTVEEHTAIGCNHKLHLILKELQFITARMKADDEAELIGDKWFAAMVVDRLCLI 471
QY 474 IFTLFTIIATLAVLLSAPHIMV 495
Db 472 VFTLFTIIATVLLSAPHIIV 493

```

RESULT 9

```

Q8T7S3
ID Q8T7S3 PRELIMINARY; PRT; 494 AA.
AC Q8T7S3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nicotinic acetylcholine receptor Dalpha6 subunit variant type I.
GN NACRALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.";
RL Genetics 160:1519-1533(2002).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321445; AAML3392.1; -.
DR FlyBase; FBgn0032151; nacr-alpha-30D.
DR GO; GO:0016020; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neur_chan_memb.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.

```

```

DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 494 AA; 56095 MW; B46BEDA63A92942 CRC64;

Query Match 67.2%; Score 1790.5; DB 5; Length 494;
Best Local Similarity 67.9%; Pred. No. 1.6e-165;
Matches 341; Conservative 53; Mismatches 83; Indels 25; Gaps 5;

QY 8 SHLAAPAGLLLLCLLWPRGARC-GYHEKRLHLHLHDHYNVLRPVVNESDPLQLSFGLT 66
Db 3 SPLPASLSFLVLLIFLAIIKESQCGPHEKRLNHLSTNTLPRPVANESPELEVFGLT 62
QY 67 LMQIIDVDEKQNLITNWLKLEWMDNLRWNTSDFGVKKDLRVP 111
Db 63 LQOIIDVDEKQNLITNWLKLEWMDNLRWNTSDFGVKKDLRVP 111
QY 127 DEGFSTYPTNVVRNGSCLYVPPGIFKSTCKIDITWFPDDQRCMKFGSWTYDGYQL 186
Db 123 DEGFDTYHTNIVVKHNGSCLYVPPGIFKSTCKIDITWFPDDQRCMKFGSWTYDGNQL 182
QY 187 DLQLODEGGDISFVNGEWELIGVPGKNEIYVNCBPPIYDITFAVVIRKTLIYFF 246
Db 183 DLVNSDEGDLSDFTNGEWELIYVNCBPPIYDITFAVVIRKTLIYFF 242
QY 247 NLIVPCVLIASMALLGFTLPDPSGKLSLGVTLISLTVPFLNVAETLQVSDAIFLGT 306
Db 243 NLIVPCVLISWALLGFTLPDPSGKLSLGVTLISLTVPFLNVAETLQVSDAIFLGT 302
QY 307 YFNCIMFVASSVSTILINLHHRHADTHEMSDWIRCVFLYVLPVWLRMSRPGSATTP 366
Db 303 YFNCIMFVASSVSTILINLHHRHADTHEMSDWIRCVFLYVLPVWLRMSRPGSATTP 362
QY 367 PARVP-PPDLELRSSKSLANVLDDDDFRHPQAQQPCCRY 423
Db 363 TILLSNRMELEKERSKSLANVLDDDDFRHTIS 411
QY 424 -----HSCFGVDY-ELSLILKEIRVITDQMRKDEDEDADISRDKFAAMVVDRLCLT 473
Db 412 GRPTTVEEHHTAIGCNHKLHLILKELQFITARMKADDEAELIGDKWFAAMVVDRLCLT 471
QY 474 IFTLFTIATLAVLLSAPHIMV 495
Db 472 VFTLFTIATVLLSAPHIIV 493

RESULT 10
Q8T7S0 PRELIMINARY; PRT; 509 AA.
AC Q8T7S0;
DT 01-JUN-2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Nicotinic acetylcholine receptor Dalpha6 subunit variant type IV.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.";
RL Genetics 160:1519-1533 (2002).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321448; AAM13395.1; -.
DR FlyBase; FBgn0032151; nAcR-alpha-30D.

```

```

DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti...; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neu_Channel_memb.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NEURCHANNEL.
DR TIGRfam; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 509 AA; 57887 MW; BE8D8E0198B0C2BD CRC64;

Query Match 66.6%; Score 1774; DB 5; Length 509;
Best Local Similarity 66.0%; Pred. No. 6.9e-164;
Matches 341; Conservative 54; Mismatches 82; Indels 40; Gaps 6;

QY 8 SHLAAPAGLLLLCLLWPRGARC-GYHEKRLHLHLHDHYNVLRPVVNESDPLQLSFGLT 66
Db 3 SPLPASLSFLVLLIFLAIIKESQCGPHEKRLNHLSTNTLPRPVANESPELEVFGLT 62
QY 67 LMQIIDV-----DEKQLITNWLKLEWMDNLRWNTSDFGVKKDLRVP 111
Db 63 LQOIIDVDEKQNLITNWLKLEWMDNLRWNTSDFGVKKDLRVP 111
QY 112 PHRLKPDVLMNSADEGDSYPTNVVRNGSCLYVPPGIFKSTCKIDITWFPDDQRC 171
Db 123 PNKLKPDVLMNSADEGDSYPTNVVRNGSCLYVPPGIFKSTCKIDITWFPDDQRC 182
QY 172 CEMKFGSWTYDGYQLDLOLQDEGGDISFVNGEWELIGVPGKNEIYVNCBPPIYDI 231
Db 183 CEMKFGSWTYDGNQLDLVNSDEGDLSDFTNGEWELIYVNCBPPIYDI 242
QY 232 TFAVVIRKTLIYFFNLIVPCVLIASMALLGFTLPDPSGKLSLGVTLISLTVPFLNVA 291
Db 243 TFTIQRRTTLYFFNLIVPCVLISWALLGFTLPDPSGKLSLGVTLISLTVPFLNVA 302
QY 292 ETWPATSDAVPLIGYFNCIMFVASSVSTILINLHHRHADTHEMSDWIRCVFLYVLP 351
Db 303 ETLQVSDAIFLIGYFNCIMFVASSVSTILINLHHRHADTHEMSDWIRCVFLYVLP 362
QY 352 WVLRRSRPGSATTPPARVP-PPDLELRSSKSLANVLDDDDFRHPQAQQPCCRY 410
Db 363 WILRMGRFGKTRKTRILLSNRMELEKERSKSLANVLDDDDFRHTIS 414
QY 411 YRGGEENGAGLAA-----HSCFGVDY-ELSLILKEIRVITDQMRKDEDEDADISR 458
Db 415 ---GSQTAIGSSASGRPTTVEEHHTAIGCNHKLHLILKELQFITARMKADDEAELIG 471
QY 459 DKKFAAMVVDRLCLIFLFTIATLAVLLSAPHIMV 495
Db 472 DKKFAAMVVDRLCLIFLFTIATVLLSAPHIIV 508

RESULT 11
Q8T7R9 PRELIMINARY; PRT; 523 AA.
ID Q8T7R9;
AC Q8T7R9;
DT 01-JUN-2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Nicotinic acetylcholine receptor Dalpha6 subunit variant type V.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

```

```

RN  SEQUENCE FROM N.A.
RP  MEDLINE=21969411; PubMed=11973307;
RA  Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT  "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT  Dalphas5, Dalphas6 and Dalphas7, in Drosophila melanogaster Identify a
RT  New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT  Mediated A-to-I Pre-mRNA Editing.";
RL  Genetics 160:1519-1533(2002)
CC  -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC  -|- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR  EMBL; AF321449; AM13396.1; -
DR  FlyBase; FBgn0032151; nAChR-alpha-30D.
DR  GO; GO:0016021; C:integral to membrane; IEA.
DR  GO; GO:0005230; F:intracellular ligand-gated ion channel activity; IEA.
DR  GO; GO:0005216; F:ion channel activity; IEA.
DR  GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR  GO; GO:0006811; P:ion transport; IEA.
DR  GO; GO:0007268; P:synaptic transmission; IEA.
DR  InterPro; IPR006201; Neur channel.
DR  InterPro; IPR006202; Neur channel.
DR  InterPro; IPR006029; Neur channel_memb.
DR  Pfam; PF02931; Neur_chan_LBD; 1.
DR  Pfam; PF02932; Neur_chan_memb; 1.
DR  PRINTS; PR00252; NRIONCHANNEL.
DR  TIGRfams; TIGR00860; LIC; 1.
DR  PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
KW  Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW  Transmembrane.
SQ  SEQUENCE 523 AA; 59110 MW; 1C200AF4F87F841 CRC64;

Query Match          66.4%; Score 1769; DB 5; Length 523;
Best Local Similarity 64.4%; Pred. No. 2.2e-163;
Matches 342; Conservative 53; Mismatches 82; Indels 54; Gaps 6;

Qy  8 SHLAAPAGLLLLCLLWPGARC-GYHEKLLHLLDHYNVLPRPVNESDPLQSPGLT 66
Db  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy  3 SPLPASLSFLVLIIFLAIKESQCGPHEKLLHLLSYNTLPRPVANSEPLEVKFGLT 62
Db  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy  67 LMOIIDVDEKNOLLITNWLKLENDNNLFWNTSDFGVKDLRPPHRLKPKDVLVYN 126
Db  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy  63 LQGLIIDVDEKNOLLITNWLKLENDNNLFWNTSDFGVKDLRPPHRLKPKDVLVYN 122
Db  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy  127 DEGFDSYTPNVVNRNGSLYVPPGIFKSTCKIDITWFFDDQRCMKFGSWTYDGYQL 186
Db  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy  123 DEGFDTGTYHTNIVKNGSLYVPPGIFKSTCKIDITWFFDDQRCMKFGSWTYDGNQL 182
Db  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy  187 DLQIQEGGDISFVTVNGEWELGVPGKNEIYNNCCPEPYDITFAVIRKTLTYFF 246
Db  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy  183 DLVLNSDGDLSDFITNGEWYLLAMPKKNITVYACCPFPYVDITFIQIRRTLYFF 242
Db  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy  247 NLIVPCVLIASMLLGLTLPDPSGKLSLGVTLISLTFLNVAETMPATSDAV---- 301
Db  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy  243 NLIVPCVLISSMLLGLTLPDPSGKLTGLVTLLSLTVLNLVAESNPTTSDAVPLIG 302
Db  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy  302 -----PLLGTYFNCIMFMVASSVSTLILNHYHHRHADTHE 337
Db  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy  303 TILLSLTVLNLVAETLPQVSDAIFLLGTYFNCIMFMVASSVLTVVVNLVHRTADIE 362
Db  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy  338 MSWIRCVFLYLPWLVRMRGSGATTPPARVP-PPDLELRSSKSLIANVLIDDDD 396
Db  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy  363 MPPWIKSVFLQPLWILRMGRPGKTRKRTILLNRMKELEKRSKSLIANVLIDDDD 422
Db  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy  397 FRHPQAQPCCRVYRGEENGAGLAA-----HSCFGVDY-ELSLILKEIRVIT 444
Db  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy  423 FRHTIS-----GSQTAGSSASFRPTVEHHTAIGCNKDLHLILKELOFIT 471
Db  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy  445 DQMRKDEDDADISRWKFAAMVDRCLLIIFTFTIATIALVLSAPHIMV 495
Db  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy  472 ARNRKADDEAEELIGDWKFAAMVDRFCLIVFTFTIATVTLVLSAPHIIV 522
Db  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 12

```

Q9VL79
ID  Q9VL79; PRELIMINARY; PRT; 554 AA.
AC  Q9VL79;
DT  01-MAY-2000 (TrEMBLrel. 13, Created)
DT  01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT  01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE  CG4128 protein.
GN  NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS  Drosophila melanogaster (Fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC  Ephydroidea; Drosophilidae; Drosophila.
OX  NCBI_TaxID=7227;
[1]
RN  SEQUENCE FROM N.A.
RP  STRAIN=Berkely;
RC  MEDLINE=20196006; PubMed=10731132;
RX  Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA  Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA  George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA  Sutton R.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA  Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA  Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA  Abril J.F., Agbayani A., An H.-U., Andrews-Frankkoch C., Baldwin D.,
RA  Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA  Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA  Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA  Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA  Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA  de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA  Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA  Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA  Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA  Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA  Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA  Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA  Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA  Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA  Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA  Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA  Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA  Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA  Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA  Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA  Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA  Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA  Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA  Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA  Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA  Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA  Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA  Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA  Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT  "The genome sequence of Drosophila melanogaster.";
RL  Science 287:2185-2195(2000).
[2]
RN  SEQUENCE FROM N.A.
RP  Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA  Evans C.A., Gocayne J.D., Ananides P.G., Brandon R.C., Rogers Y.,
RA  Banon J., An H., Baldwin D., Banon J., Beeson K.Y., Busam D.A.,
RA  Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA  Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA  Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,
RA  Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA  Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA  McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA  Pacleb J., Fargas V., Park S., Patel S., Pfeiffer B.,
RA  Phouenanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA  Scapleton M., Strong R., Swirskas R., Tector C., Tyler D.,
RA  Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT  "Sequencing of Drosophila melanogaster genome.";
RL  Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]

```


QY 375 DLRLRRSSKLLANVLIDDDP-----HPQAQPO-----CCRYRGGREN-----CAGL 421
Db 370 DLNISPVGASTNGNLIIYI--GFGMTIHH-YATSPDSGWICSLVATGEDVLLPQAQA 426
QY 422 AAHSCFGV-DYELSLILKEIRVITDQMRKDEDADISRDWKFAAMVVDRLCLIIFTLFTI 480
Db 427 SSVSSSGFGETELSKILDEVRIYSKRFDQDEEDTVCNEWKFAASVIDRLCLMAFSLFTI 486
QY 481 IATLAVLLSAPHIM 494
Db 487 LCTIGILMSAPNFV 500

RESULT 15

Q7T2U0 PRELIMINARY; PRT; 555 AA.
AC Q7T2U0
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Nicotinic acetylcholine receptor alpha 7b subunit (Fragment).
GN A7B.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RA Jones A.K., Elgar G., Sattelle D.B.;
RT "The nicotinic acetylcholine receptor gene family of the pufferfish,
RT Fugu rubripes";
RL Genomics 0:0-0(2003).
DR EMBL; AY298752; AAP57216.1; -.
KW Receptor.
FT NON_TER
SQ SEQUENCE 555 AA; 61010 MW; B042A3E4430A2B7C CRC64;

Query Match 47.0%; Score 1253; DB 13; Length 555;
Best Local Similarity 44.7%; Pred. No. 4.7e-113;
Matches 246; Conservative 86; Mismatches 126; Indels 92; Gaps 9;
QY 31 GYHEKRLHLLHLYNVRPVNESDPLQLSGLTMOIIDVDEKNQLIINIKLEW 90
Db 4 GPEQRTLLKXLLKDYNRMRPVANDSQPLTVVFTLSLIQIMDVDEKNQLITNIMLRMSG 63
QY 91 NDMLRWNTSDFGVKDLRVPVHRLMKPDVLMYNSADEGFDSTPTNVVVRNNGSLYVP 150
Db 64 FDHLYQWQSEHPGVKNLFTTDQIWTDPILLYNSADDDFDSTFTKNVLVNSGYAEYQP 123
QY 151 PGIFKSTCIDITWPFDDQRCMKFGSWTYDGYQLDQLQDEGGDISFVYNGEWELI 210
Db 124 PGIFMSTCNVDVWPFDFIQRCLEKFGSWTYDGLLDLQWNE---ADISGYMANGEWDLI 180
QY 211 GVPGRKREIYVNCPEPIDITFAVVRKTYLYFFENLIVPCVLIASMALLGFTLPDPSG 270
Db 181 GVPGRTRNEVYOCCKEPPYPAVTFVVAIRRTIYALNLLIPCVLLSSMTLLIFVLPSDG 240
QY 271 EKLSLGVTLISLTVFLNVAETMPATSDAVPLGLTYFNCIMFMVASSVSVSTILINLHH 330
Db 241 EKISLITVLLSITVFLVAETMPATSDVPLIGQYFASIMIIYVMSVVAIVVVLQYHH 300
QY 331 RHADTHEMSDWIRCVFLYLPWVLRMSRFGSATTP--PP-----ARVPPPP 374
Db 301 HDPNGGNMFKWQLVQLLOWVANFLMRKRGKXDSRPPCAPHLRRCSSGSGSIPNPP 360
QY 375 DLRL-----RERSKSLIANVL-----DIDDFRHPQAQ-- 404
Db 361 DHALHPLHPQGLAPLQPGHLAGQPHVHAQSSANNNGNLVYLFQSVDSAGLPEPIQRN 420
QY 405 -----PQCCRY-----YRGGEENGAGLAHS 425

Db 421 NISTGPPRVAGSPPHLPSPQFCSSPPPPASNMMDTGCPSSTVSSGGGFGG---GGGLGGCS 477
QY 426 CFGV-DYELSLILKEIRVITDQMRKDEDADISRDWKFAAMVVDRLCLIIFTLFTIATL 484
Db 478 ASAVGDPQLHALLLEEVRFVADRFREQDEVSGAADQWKFAAGVIDRLCLVAFSVFNIICTI 537
QY 485 AVLLSAPHIM 494
Db 538 SILMSAPNFV 547

Search completed: May 7, 2004, 11:38:38
Job time : 33.1579 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2004, 11:50:34 ; Search time 3463.3 Seconds
(without alignments)
4276.744 Million cell updates/sec

Title: US-09-303-232-4

Perfect score: 2665

Sequence: 1 MCGRRARRSHAPAGLLLL.....LFTIATLAVLLSAPHIMVS 496

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-O=/cn2_1/USPTO.spool/US030232/runat_07052004.101110.23893/app_query.fasta_1.2261
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0303232 @CN 1.1 4577 @runat_07052004.101110.23893 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_fod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
C					
1	53.8	1433	885	12	BG632919 GH16126.3
2	36.0	960	2296	14	CD013901
3	35.4	942.5	1436	29	AY402873 Homo sapi
4	34.8	928	1864	11	AK053497 Mus muscu
5	34.8	928	2916	11	AK051730 Mus muscu
6	34.8	928	3126	11	AK080415 Mus muscu
7	34.8	927.5	1201	9	AL530299 AL530299
8	34.3	914.5	4037	11	AK049722 Mus muscu
9	34.3	914.5	4046	11	AK051742 Mus muscu
10	34.3	913.5	1436	29	AY402875 Mus muscu
11	34.0	905	1454	29	AY402878 Mus muscu
12	33.6	895.5	1442	29	AY402876 Mus muscu
13	33.5	892.5	922	13	BUS15857
14	33.4	889	1374	29	AY406230
15	33.4	889	1374	29	AY406232
16	33.4	889	4290	11	AK029177 Mus muscu
17	33.1	881.5	2010	11	AK080475 Mus muscu
18	32.9	877.5	1442	29	AY402877 Mus muscu
19	32.6	869	2940	11	AK034228
20	32.6	869	3230	11	AK083157
21	32.5	866.5	1781	14	CD013889
22	32.0	853	4589	11	AK030464 Mus muscu
23	31.7	845.5	1374	29	AY406231 Pan trogl
24	31.4	835.5	3483	11	AK081254 Mus muscu
25	31.2	832.5	1436	29	AY402874
26	30.1	801.5	1751	14	CD013905
27	30.1	801.5	2110	14	CD013902
28	30.1	801.5	2110	14	CD013903
29	29.2	779	1466	14	CD013891
30	29.2	777	2074	11	AK080554 Mus muscu
31	28.9	771.5	2257	11	AK017571 Mus muscu
32	28.9	771	2513	11	AK033068 Mus muscu
33	28.9	769	1034	13	BX403124 BX403124
34	28.6	762.5	908	13	BUI49265
35	27.5	733.5	607	9	AI292581 GH1518.5
36	27.0	720.5	1596	14	CD013888
37	26.8	715.5	615	14	CB149460
38	26.4	704	1603	29	AY411327 Homo sapi
39	26.3	700.5	2948	11	AK081344 Mus muscu
40	25.8	686.5	833	14	CB245337 UI-M-FY0-
41	25.6	682.5	1101	29	CNS0001F
42	25.4	677.5	2782	11	AK047974
43	25.4	676.5	1007	29	CNS000HU
44	25.3	674.5	902	29	AY407186 Mus muscu
45	25.2	672	902	29	AY407184 Homo sapi

ALIGNMENTS

RESULT 1
BG632919/c
LOCUS
DEFINITION
BG632919
885 bp mRNA linear EST 23-APR-2001
GH16126.3prime GH Drosophila melanogaster head p072 Drosophila
melanogaster cDNA clone GH16126.3 similar to CG4128: FBan0004128
'ion channel' located on: 2L 30D1-30E1; 04/10/2001, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BG632919.1 GI:13758409
EST.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 885)
 Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
 Lewis, S. and Rubin, G.M.
 BDGP/HMI Drosophila EST Project
 Unpublished (2001)
 Other_ESTs: GH16126.5prime
 Contact: Stapleton, M.
 BDGP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
 Based upon the presence of a XhoI site followed by a run of 14 or
 more T residues at the beginning of the sequence, this clone was
 polyadenylated. The resulting Poly-T sequence has been removed. hit
 genomic AB003511: armX [18792641,19136447]
 estimated-cyto:18A3-18C6: 04/10/2001
 Plate: GH.161 row: C column: 2
 High quality sequence stop: 784.
 Location/Qualifiers
 1..885

FEATURES

Location/Qualifiers
 1..885
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /clone="GH16126"
 /sex="male and female"
 /dev_stage="adult"
 /lab_host="DH5 - alpha"
 /clone_lib="GH Drosophila melanogaster head pOT2"
 /note="Organ: head; Vector: pOT2; Site 1: EcoRI; Site 2:
 XhoI; Sized fractionated cDNAs were directly ligated into
 pOT2. Plasmid cDNA library."

ORIGIN

Alignment Scores:
 Pred. No.: 5,12e-122 Length: 885
 Score: 1433.00 Matches: 260
 Percent Similarity: 97.18% Conservativity: 16
 Best Local Similarity: 91.55% Mismatches: 6
 Query Match: 53.77% Indels: 2
 DB: 12 Gaps: 1

US-09-303-232-4 (1-496) x BG632919 (1-885)

65 LeuThrLeuMetGlnIleAspValAspGluLysAsnGlnLeuLeuIleThrAsnIle 84
 |||||
 885 CTAACACTCATGCAGATTATCGATGTGCACGAGAGAATCAACTGCTTATAACGAATATT 826
 |||||
 85 TrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspPheGlyGly 104
 |||||
 825 TGGCTCAAAATGGAAATGGAATGAACGATGATGAATCTTCATGGAATTCGATGAGTTTCGGTGGT 766
 |||||
 105 VallysAspLeuArgValProHisArgLeuTrpLysProAspValLeuMetTyrAsn 124
 |||||
 765 GTGCGGATCTGCGAATTCGCCCATCGCTATGGAACCGGATGACTGATGTACAAAC 706
 |||||
 125 SerAlaAspGluGlyPheAspSerThrTyrProThrAsnValValArgAsnGly 144
 |||||
 705 AGTCCGACGAGGGCTTCGATGGACCGTACCCCAAAATGGTGGTTCCGAAATAATGGG 646
 |||||
 145 SerCysLeuTyrValProGlyIlePheLysSerThrCysLysIleAspIleThrTrp 164
 |||||
 645 AGCTGCTGCTACGTACCCGAGGTATATTAAAGTCAAGCTGTAGATCGACATTACGGTGG 586
 |||||
 165 PheProPheAspAspGlnArgCysGluMetLysPheGlySerTrpThrTyrAspGlyTyr 184
 |||||
 595 TITTCATTTCGACGATCAGAGATGTAATGAAATTTGTTGCTGGACCTACGATGGGTTT 526
 |||||
 185 GlnLeuAspLeuGlnLeuGlnAspGluGlyGlyAspIleSerSerPheValThrAsn 204
 |||||
 525 CAGTTGGACCTGACGTTCGAGACGACGAGCTGGTGGCGACATTCTTAGCTTTATAACCAAT 466
 |||||

205 GlyGluTrpGluLeuIleGlyValProGlyLysArgAsnGluIleTyrTyrAsnCysCys 224
 |||||
 465 GCGGAATGGGACTTGTAGTGTGCCGCTAAACGAAATCAATCTACTATAATTCCTGC 406
 |||||
 225 ProGluProTyrIleAspIleThrPheAlaValAlaIleArgArgLysThrLeuTyrTyr 244
 |||||
 405 CCAGAACCTTATATTGACATAACATTGCGCATTTTGTAAAGCGCAAAAGCTTGTACTAT 346
 |||||
 245 PhePheAsnLeuIleValProCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThr 264
 |||||
 345 TTTTTCAAATCTGATTGTGCGGTGCGTACTGATCGCTCCATGCGACTGCTAGGTTTACA 286
 |||||
 265 LeuProProAspSerGlyGluLysLeuSerLeuGlyValThrIleLeuLeuSerLeuThr 284
 |||||
 285 CTGCCACCAAGATTCTCGTGAAAGCTTTCGTTGGAGTTTACAAATTTATATCGCTTACA 226
 |||||
 285 ValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValProLeuLeu 304
 |||||
 225 GTCTTCTCAACATGTGTGCGGAAACAATGCCGCGACCTCCGATGGGTACCGCTGCTC 166
 |||||
 305 GlyThrTyrPheAsnCysIleMetPheMetValAlaSerSerValValSerThrIleLeu 324
 |||||
 165 GGAACCTATTTCATTTGCATTATGTTTATGGTGGCTCATCAGTTGTGTCAACATACTT 106
 |||||
 325 IleLeuAsnTyrHisArgHisAlaAspThrHisGluMetSerAspTrpIleArgCys 344
 |||||
 105 GTCTCAATTTATCATCATAGAAATCCAGATACGATGAAATGATGATGAAATGGGTA----- 52
 |||||
 345 ValPheLeuTyr 348
 |||||
 51 ATAGTTTATAT 40

RESULT 2

CD013901 2296 bp mRNA linear EST 21-OCT-2003
 LOCUS 90134548 Single gene library Homo sapiens cDNA, mRNA sequence.
 DEFINITION CD013901
 ACCESSION CD013901
 VERSION CD013901.1 GI:37777431
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 2296)
 Au-Yang, J., Fu, G.K., Wilson, A.D., Yang, J., Chien, D., Hawkins, P.R.,
 PCR isolation and cloning of novel splice variant mRNAs from known
 drug target genes
 Unpublished (2003)

JOURNAL

Incyte Corporation

Contact: Jin, P.

3160 Porter Drive, Palo Alto, CA 94304, USA

Tel: 650 621 8639

Fax: 650 621 8965

Email: pjin@incyte.com

FEATURES

Location/Qualifiers

source

1..2296
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="Single gene library"
 /note="Vector: pDrive Cloning Vector; RT-PCR was performed
 using gene-specific primers flanking the open-reading
 frame. PCR products were subcloned into pDrive Cloning
 Vector and sequenced completely using M13 forward and
 reverse primers. Sequencing gaps were closed by
 re-sequencing using primers flanking the gapped areas."

ORIGIN

Alignment Scores:
 Pred. No.: 1.31e-77 Length: 2296
 Score: 960.00 Matches: 199
 Percent Similarity: 56.45% Conservativity: 90

[illegible]

```

gene      <1..>1436
          /gene="CHRNA3"
          /locus_tag="HCM1369"

ORIGIN
Alignment Scores:
Pred. No.: 2,75e-76      Length: 1436
Score: 942.50           Matches: 194
Percent Similarity: 58.25% Conservative: 92
Best Local Similarity: 39.51% Mismatches: 153
Query Match: 35.37%      Indels: 53
DB: 29                  Gaps: 6

US-09-303-232-4 (1-496) x AY402873 (1-1436)
Qy 28 AlaArgCysGlyTyrHisGluLysArgLeuLeuHisHisLeuLeuAspHisTyrAsnVal 47
Db 3 GCCAGGCGCTCAGAGCGTGAGCACCGTCTATTGAGCGGCTGTTTGAAGATTACAATGAG 62
Qy 48 LeuGluArgProValValAsnGluSerAspProLeuGlnLeuSerPheGlyLeuLeu 67
Db 63 ATATCCGGCTGTGGCCACAGTGTCTGACCCAGTCAATCCATTTGAGGTTGCTCATG 122
Qy 68 MetGlnIleLeuAspValAspGluLysAsnGlnLeuLeuLeuLeuLeuLeuLeuLeu 87
Db 123 TCTCAGCTGTGAAGGTGGATGAAGTAAACACAGATCATGGAGACCAACCTGTGCTCAAG 182
Qy 88 LeuGluTrpAsnAspMetAsnLeuArgTyrAsnThrSerAspPheGlyValLeuLeu 107
Db 183 CAATCTCGAATGACTACAGCTGAAGTGAACCCCTCTGACTATGTGGGGCAGAGTTC 242
Qy 108 LeuArgValProHisArgLeuTyrLysProAspValLeuMetTyrAsnSerAlaAsp 127
Db 243 ATCGGTCTCCTGCACAGAGATCTGGAAGCCAGACATTTGCTGTTATACAAATGCTGT 302
Qy 128 GluGlyPheAspSerThrTyrProThrAsnValValValArgAsnAsnGlySerCysLeu 147
Db 303 GGGGATTTCCAGGTGGGAGCAACAGCAAGCCCTTACTCAAGTACACTGGGAGGTGACT 362
Qy 148 TyrValProProGlyIlePheLysSerThrCysLysIleAspIleThrTrpPheProPhe 167
Db 363 TGTATCTCCGGCCATCTTTAAGAGCTCTCTGTAATAATCGACGTGACTTCCCGTTT 422
Qy 168 AspAspGlnArgCysGluMetLysPheGlySerThrTyrAspGlyTyrGlnLeuAsp 187
Db 423 GATTACCAAACTGATACCATGAAGTTCGGTTCCTGCTCCTACGATAGGCGCAATCGAT 482
Qy 188 LeuGlnLeuAspGluGlyGlyGlyAspIleSerSerPheValThrAsnGlyGluTrp 207
Db 483 CTGTCTCTGATCGGC---TCTTCATGAACCTCAAGGACTATTGGGAGAGCGCGAGTGG 539
Qy 208 GluLeuIleGlyValProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluPro 227
Db 540 GCCATCATCAAGCCCGGCTCAAAACAGACATCAAGTACAACTGCTCGAGGAGTAC 599
Qy 228 TyrIleAspIleThrPheAlaValIleArgLysThrLeuTyrTyrPhePheAsn 247
Db 500 TACCCGACATCATACTGCTGTATCATCGGGCGCTGCTGCTTGTCTACACCATCAAC 659
Qy 248 LeuIleValProCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProPro 267
Db 660 CTCATCATCCCTGCTGCTCATCTCTCTCTCACTGCTGCTGCTTCTACCTGCCCTCC 719
Qy 268 AspSerGlyGluLysLeuSerLeuGlyValThrIleLeuLeuSerLeuValPheLeu 287
Db 720 GACTGCGGTGAGAGGTGACCTGTGATTTCTGCTCTCTCTCCCTGACCGGTGTTCTC 779
Qy 288 AsnMetValAlaGluThrMetProAlaThrSerAspAlaValProLeuLeuGlyThrTyr 307
Db 780 CTGTGTGATCACTGAGACCATCCCTCCACCTCGTGTGTCATCCCTGATTGGAGAGTAC 839
Qy 308 PheAsnCysIleMetPheMetValAlaSerSerValValSerThrIleLeuIleLeuAsn 327

```

```

Db 840 CTCCTGTTCCACCATGATTTTTGTAAACCTTGTCCATCGTCATCCGCTTCTGCTCAAC 899
Qy 328 TyrHisHisArgHisAlaAspThrHisGluMetSerAspTrpIleArgCysValPheLeu 347
Db 900 GTGCACCTACAGAACCCCGACGACACACAAATCCCTCATGGGTGAAGACTGTTATTTTG 959
Qy 348 TyrTrpLeuProTrpValLeuArgMetSerArgPro-----GlySerAlaThr 363
Db 960 AACCTGCTCCAGGGTCACTGTTTCATGACCCAGCCCAACAGCAACGAGGCGCAACGCTCAG 1019
Qy 364 ThrPro-ProProAlaArgValProProProProProProProProProProProProPro 383
Db 1020 AAGCCGAGGCCCTCTACGGTGCCG-----AGCTC 1049
Qy 383 rLysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHisProGlnAlaG1 403
Db 1050 TCAAAATCTGAATTCCTTCAGCC-----GGCA 1076
Qy 403 nGlnProGlnCysCysArgTyrTyrArgGlyGlyGluGluAsnGlyAlaGlyLeuAla1 423
Db 1077 GAGTCCAAAGGC-TGCAAGGAGGGCTACCCCTGCCAGGACGGGATGTGTGTACTGCCA 1135
Qy 423 aHis-----SerCysPh 427
Db 1136 CCACCGCAGGATAAAATCTCAATTTTCAGTGTCTAACCTCAGGAGAGCTCTAGTTCTGA 1195
Qy 427 eGlyValAspTyrGluLeuSerLeu-----IleLeuLy 438
Db 1196 ATCTGTGATGCTGTGCTGCTGCTCTCTCTGCTTTGTCCACGAAATCAAGAGCATCTCCA 1255
Qy 438 sGluIleArgValIleThrAspGlnMetArgLysAspAspGluAspAlaAspIleSerAr 458
Db 1256 AAGTGTCAAGTATATGCTGAAATATGAAGACACAAATGAAGCAAGAGATTCAAGA 1315
Qy 458 gAspTrpLysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrLeuPh 478
Db 1316 TGATTGAAGATGTTGCCATGTGTGATGATGATTTCTGTTGGGTTTTCACCTGGT 1375
Qy 478 eThrIleAlaThrLeuAlaValLeuLeu 488
Db 1376 GTGCATTTAGGGACAGCAGGATTGTTTCG 1406

RESULT 4
AK053497
LOCUS
DEFINITION
Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone E130103E14 product: NEURONAL NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT, full insert sequence.
ACCESSION
VERSION AK053497.1 GI:26343494
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE
1
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
AUTHORS
2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
AUTHORS
3
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
AUTHORS
3
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
AUTHORS

```


QY	273	LeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAenMetValAlaGlu	292			
Db	892	GTACGCTGTCATCTCGTGCTCTCTCCCTGACGGTCTTCTCTCGTGATCACCGAG	951			
QY	293	ThrMetProAlaThrSerAspAlaValProLeuLeuGlyThrTyrPheAenCysIleMet	312			
Db	952	ACCATCCCTTCCACCTCACTGGTGCATCCCTTGATCGGGAGTACCTCTCTTCACCTATG	1011			
QY	313	PheMetValAlaSerSerValSerThrIleLeuLeuLeuAenTyHisArgHis	332			
Db	1012	ATTTTTCACCTTGTCCATCGTCATCACAGTCTTTGTGTCACAGTGCCAGAACT	1071			
QY	333	AlaAspThrHisGluMetSerAspTrrPileArgCysValPheLeuTyTrpLeuProTrrP	352			
Db	1072	CCGACCACACACAGATGCCACTTGGTCAAGCTGTGTTTGAACCTTCTCCCCAGG	1131			
QY	353	ValLeuArgMetSerArgProGlySerAlaThrThrProProAlaArgValProPro	372			
Db	1132	GTTCATGTTTATGACTAGGCCA-----ACCAGCACCGAGGAGAGCGCCCAAG	1179			
QY	373	ProProAspLeuGluLeuArgGluArgSerSerLysSerLeuLeuAlaAenValLeuAsp	392			
Db	1180	ACGAGGAACCTTACGGTGTGCTGACTCTCAAACTGAACTGCTTCAGCCGTGCAGACTCC	1239			
QY	393	IleAspAspAspPheArgHisProGlnAlaGlnGlnProGlnCys-----CysArgTyr	410			
Db	1240	AAAAGCTCAAGGAGGCTACCCC---TGCCAAGATGGGACATGTGGCTATTGCCACGAC	1296			
QY	411	TyrArgGlyGlyGluGluAen---GlyAlaGlyLeuAla---AlaHisSerCysPheGly	428			
Db	1297	CGTAGGGTAAAAATCTCAAATTTTCAGTGCACCACTCAAGAAGCTCCAGTTCGTGAGTCT	1356			
QY	429	ValAspTyrGluLeuSerLeu-----IleLeuLysGlu	439			
Db	1357	GTTCATGCTGTGTGTGCTCTCTCTGTCTGTCTACACAGAAATCAAGAAGCCATCCAAAGT	1416			
QY	440	IleArgValIleThrAspGlnMetArgLysAspAspGluAspAlaAspIleSerArgAsp	459			
Db	1417	GTGAAGTATATTGCTGAAATATGAAGACACAGAATGAGCAAGAGATTCAAGATGAT	1476			
QY	460	TrpLysPheAlaAlaMetValValAspArgLeuLeuLeuIlePheThrLeuPheThr	479			
Db	1477	TGGAAGTATGTGTGCATGTGATGATGATTTTCTCGGTTCATCTCGTGTGTC	1536			
QY	480	IleIleAlaThrLeuAlaValLeuLeu	488			
Db	1537	ATTTAGGAACATGCAGGATTAATTTCTG	1563			
RESULT 5						
AK051730						
LOCUS	AK051730	2916 bp	mRNA			
DEFINITION	Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length enriched library, clone:D130068A06 product:NEURONAL NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT, full insert sequence.					
ACCESSION	AK051730					
VERSION	AK051730.1	GI:26342155				
KEYWORDS	HTC; CAP trapper.					
SOURCE	Mus musculus (house mouse)					
ORGANISM	Mus musculus					
REFERENCE	<p>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.</p>					
AUTHORS	1 Carninci, P. and Hayashizaki, Y.					
TITLE	High-efficiency full-length cDNA cloning					
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)					
MEDLINE	99279253					
PUBMED	10349636					
REFERENCE	2					
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.					
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes					
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)					
MEDLINE	20499374					
PUBMED	11042159					
REFERENCE	3					
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.					
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer					
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)					
MEDLINE	20530913					
PUBMED	11076861					
REFERENCE	4					
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.					
TITLE	Functional annotation of a full-length mouse cDNA collection					
JOURNAL	Nature 409, 685-690 (2001)					
REFERENCE	5					
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.					
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs					
JOURNAL	Nature 420, 563-573 (2002)					
REFERENCE	6 (bases 1 to 2916)					
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saigo, R., Saichou, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.					
TITLE	Direct Submission					
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)					
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ Location/Qualifiers					
FEATURES	source					
1. .2916						
/organism="Mus musculus"						
/mol_type="mRNA"						
/strain="C57BL/6J"						
/db_xref="FANTOM_DB:D130068A06"						
/db_xref="MGI:2420066"						
/db_xref="taxon:10090"						
/clone="D130068A06"						
/tissue type="spinal ganglion"						
/clone lib="RIKEN full-length enriched mouse cDNA library"						
/dev_stage="12 days embryo"						
170. .1669						
/note="unnamed protein product; NEURONAL NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT (SPTR)AAL58471, evidence: FASTY, 99.8%ID, 100%length, match=1497 putative"						
/codon_start=1						
CDS						

REFERENCE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muranatsu, M. and Hayashizaki, Y.
AUTHORS		Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL		Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE		20499374
PUBMED		11042159
REFERENCE	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J., Okazaki, Y., Muranatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
AUTHORS		RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL		Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE		20530913
PUBMED		11076861
REFERENCE	4	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
AUTHORS		Functional annotation of a full-length mouse cDNA collection
JOURNAL		Nature 409, 685-690 (2001)
MEDLINE		11076861
PUBMED		11076861
REFERENCE	5	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
AUTHORS		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL		Nature 420, 563-573 (2002)
MEDLINE		12000000
PUBMED		12000000
REFERENCE	6	(bases 1 to 3126)
AUTHORS		Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saigo, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
AUTHORS		Direct Submission
JOURNAL		Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
MEDLINE		20499374
PUBMED		11042159
REFERENCE	7	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
AUTHORS		Please visit our web site for further details.
JOURNAL		URL: http://genome.gsc.riken.go.jp/
MEDLINE		20499374
PUBMED		11042159
REFERENCE	8	Location/Qualifiers
AUTHORS		1. 3126
JOURNAL		/organism="Mus musculus"
MEDLINE		/mol_type="mRNA"
PUBMED		/strain="C57BL/6J"
REFERENCE	9	/db_xref="FANTOM,DB:A730007P14"
AUTHORS		/db_xref="MGI:2407635"
JOURNAL		/db_xref="taxon:10090"
MEDLINE		/clones="A730007P14"
PUBMED		/tissue_type="cerebellum"
REFERENCE	10	/clone_lib="RIKEN full-length enriched mouse cDNA library"
AUTHORS		/dev_stage="7 days neonate"

```

QY 213 ProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThr 232
Db 785 CCGGCTACAAACATGAATCAAGTACACTGCTGTGAGAGATCTACCAAGACATCAG 844
QY 233 PheAlaValIleAlaMetValSerValSerValSerValSerValSerValSerVal 252
Db 845 TACTCGCTATACATTCGCGCGCTGCTGTCTACACCATCAACCTCATCTCCGTGC 904
QY 253 ValIleIleAlaMetMetAlaLeuLeuGluPheThrLeuProProAspSerGlyGluLys 272
Db 905 CTGCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 964
QY 273 LeuSerLeuGlyValThrIleLeuSerLeuThrValPheLeuAsnMetValAlaGlu 292
Db 965 GTGAGCTCTGCACTCTCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1024
QY 293 ThrMetProAlaThrSerAspAlaValProLeuLeuGlyThrTyrPheAsnCysIleMet 312
Db 1025 ACCATCCCTTCCACCTCACTGGTCACTCCCTTGATCGGGAGTACCTCTCTCTCTCTCT 1084
QY 313 PheMetValAlaSerSerValSerValSerValSerValSerValSerValSerVal 332
Db 1085 ATTTTGTCTACCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1144
QY 333 AlaAspThrHisGluMetSerAspTyrIleArgCysValPheLeuTyrTrpLeuProTyr 352
Db 1145 CCGACACACACAGATGCCACTTGGGTCAAGCTGTGTTTGAACCTTCTCCCGAG 1204
QY 353 ValIleuArgMetSerArgProGlySerAlaThrThrProProAlaArgValProPro 372
Db 1205 GTCATGTTTATGACTAGGCCA-----ACCAGCAGGAGGAGGAGTACCTCTCTCTCTCT 1252
QY 373 ProProAspLeuGluLeuArgGluArgSerSerLysSerLeuLeuAlaAsnValLeuAsp 392
Db 1253 ACAGGAACTTCTACCGTGTGAGCTCTCAAACTGAAGCTGCTTCAAGCTGTCAGACTCC 1312
QY 393 IleAspAspAspPheArgHisProGluAlaGlnProGlnCys-----CysArgTyr 410
Db 1313 AAAAGTGCAGGAGGAGTACCC---TGCAGATGGGACATGTGGCTATTGCGCCACC 1369
QY 411 TyrArgGlyGlyGluGluAsn---GlyAlaGlyLeuAla---AlaHisSerCysPheGly 428
Db 1370 CGTAGGTAATAATCTCAAAATTTCAAGTGCCAACTCAAGAAGCTCCAGTCTGAGTCT 1429
QY 429 ValAspTyrGluLeuSerLeu-----IleLeuLysGlu 439
Db 1430 GTTGATGCTGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1489
QY 440 IleArgValIleThrAspGlnMetArgLysAspAspGluAspAlaAspIleSerArgAsp 459
Db 1490 GTGAAGTATTGCTGAATAATGAAGACAGATGTAGCCAAAGAGATTCAGATGAT 1549
QY 460 TrpLysPheAlaAlaMetValValAspArgLysCysLeuIleIlePheThrLeuPheThr 479
Db 1550 TGGAAAGTATTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1609
QY 480 IleIleAlaThrLeuAlaValLeuLeu 488
Db 1610 ATTTTAGGAAGTCAAGATTATTTCTG 1636

RESULT 7
AL530299 1201 bp mRNA linear EST 23-MAY-2003
LOCUS
DEFINITION AL530299 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
cDNA clone CS0DD007YP05 5-PRIME, mRNA sequence.
ACCESSION AL530299
VERSION AL530299.2 GI:31068132
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)

```

```

AUTHORS Li, W. B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 13, 2001 this sequence version replaced gi:12793792.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7646.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DD007CH03QP1&cluster=7646.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DD007CH03QP1.
FEATURES
Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD007YP05"
/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 5,2e-75 Length: 1201
Score: 927.50 Matches: 175
Percent Similarity: 72.78% Conservative: 63
Best Local Similarity: 53.52% Mismatches: 78
Query Match: 34.80% Indels: 12
DB: Gaps: 2

US-09-303-232-4 (1-496) x AL530299 (1-1201)
QY 5 AlaArgSerHisLeuAlaProAlaGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuTyr 24
Db 130 GCTGCTCGCGGAGGTCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 171
QY 25 ProArgGlyAlaArgCysGlyTyrHisGluLysArgLeuLeuHisHisLeuLeuAspHis 44
Db 172 CTTGCACTGCTCCCTGCAAGCAGTTCACAGGAGGAGTTCACAGGAGGAGTTCACAGGAG 231
QY 45 TyrAsnValLeuGluArgProValValAsnGluSerAspProLeuGlnLeuSerPheGly 64
Db 232 TACAATCCCTTGAGAGAGCCCGTGGCCATGATCGCAACACCTACCGTCTTCTTCTTCC 291
QY 65 LeuThrLeuMetGlnIleIleAspValAspGluLysAsnGlnLeuLeuLeuLeuLeuLeu 84
Db 292 CTGAGCTCTCTGAGATCATGAGCTGAGTGGATGAGAGACCAAGTTTAAACCAACAT 351
QY 85 TrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspPheGlyGly 104
Db 352 TGGCTGCAATATGCTGGACAGATCACTATTACAGTGAATGCTGTCAGATATCCAGGG 411
QY 105 ValLysAspLeuArgValProProHisArgLeuTrpLysProAspValLeuMetTyrAsn 124
Db 412 GTGAAGACTGTTCTGTTCCAGATGGCCAGATTGGAAACCAACATCTTCTCTATAAC 471
QY 125 SerAlaAspGluGlyPheAspSerThrTyrProThrAsnValValValArgAsnAsnGly 144
Db 472 AGTGTGATGAGCGCTTTGACGCCACATCCACACTAACCGTGTGTGATTTCTCTGG 531
QY 145 SerCysLeuTyrValProProGlyIlePheLysSerThrCysLysIleAspIleThrTrp 164
Db 532 CATTCAGTWCCTGCTCCAGGCATATTCAAGAGTTTCTGCTACATCGATGATGACGTGG 591
QY 165 PheProPheAspAspGlnArgCysGluMetLysPheGlySerTrpThrTyrAspGlyTyr 184

```



```

Db      1082 CTGACGTACCGCTGGTGGGAAGTACCTCATGTTCCACGATGGTGTAGTACCTTCTCC 1141
QY      319 ValValSerThrIleLeuLeuLeuAsnTyrHisHisArgHisAlaAspThrHisGluMet 338
Db      1142 ATGCTCACTAGCGTGTGTGCTCAATGTGCACACCGTTCGCTACCGACGACCATG 1201
QY      339 SerAspTyrIleArgCysValPheLeuTyrTrpLeuProTrpValLeuArgMetSerArg 358
Db      1202 GCOCCTCGGTCAAGGTGGTCTTCCTGGAGAAGCTGCCACCTCTCTTCCTGCGACGAG 1261
QY      359 ProGlySerAlaThrThrProProAlaArgValProProProProProProProProPro 378
Db      1262 CCAGCCACCGCTGT-----GCAGT-----CAAGCCCTGCGCTTG 1297
QY      379 ArgGluArgSerSerIysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArg 398
Db      1298 CGAAGGCGCCAGCGGAACGTGAGGGGCGCAGGCACACTA-----TTCTTCGCG 1345
QY      399 HisProGlnMlaGlnGlnProGlnCysCysArgTyrTyrArgGlyGlyGluGluAsn--- 417
Db      1346 GAAGGTCTGTCACTGACCATGACCATGCTCTTTGTCAATCCTGCATCAATGCGAGGCTTG 1405
QY      418 ---GlyValaglyLeuAlaAlaHisSerCysPheGlyValAspTyrGluLeu----- 433
Db      1406 GCTGGGGCTTCCAGCTGAGCCGCTGCAGCCGGCTGGGGCGCTGCATGGGCCCATGC 1465
QY      434 SerLeuIleLeuIysGlu-----IleArgValIleThrAspGlnMetArgLys 449
Db      1466 AGCTGTGGCTCCGGGAAGCTGTGACGGTGTACGCTTATTGGGACCATATGCGAAGT 1525
QY      450 AspAspGluAspAlaAspIleSerArgAspTyrIysPheAlaAlaMetValValAspArg 469
Db      1526 GAAGATGATACACAGAGTGTGAGGAGGATGGAAATACCTGCCATGGTGCATCGACCGC 1585
QY      470 LeuCysLeuIleIlePheThrLeuPheThrIleAlaThrLeuAlaValLeuLeu 488
Db      1586 CTGTTCTCTGGATCTTGTCTGTGTCTGTGTGGGACCATGGCAGTTCCTG 1642

RESULT 10
LOCUS   AY402875
DEFINITION Mus musculus CHRNA3 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY402875
VERSION   AY402875.1 GI:39758858
KEYWORDS  GSS.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1436)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferrlera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1436)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferrlera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
Direct Submission
TITLE Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
JOURNAL
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
Location/Qualifiers
1..1436
source
/organism="Mus musculus"

```

```

/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1436
/gene="CHRNA3"
/locus_tag="HCML369"

ORIGIN

Alignment Scores:
Pred. No.: 1,33e-73 Length: 1436
Score: 913.50 Matches: 187
Percent Similarity: 59.19% Conservative: 90
Best Local Similarity: 39.96% Mismatches: 172
Query Match: 34.28% Indels: 19
DB: 29 Gaps: 7

US-09-303-232-4 (1-496) x AY402875 (1-1436)
QY 34 GLuLyArGlLeuLeuHisLeuLeuAspHisTyrAsnValLeuGluArgProValVal 53
Db 21 GAGCACCAGCTGTTCAGTACCTGTTTGAAGATTACAGAGATCATCCGCCGGTGGCT 80
QY 54 AsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAspVal 73
Db 81 AACGTGCTCCATCTCTGTATCATCCAGTTTGAAGTGTCTATGCTCAGCTGGTGAAGTG 140
QY 74 AspGluLyAsnGlnLeuLeuThrAsnIleTrpLeuLysLeuGluTrpAsnAspMet 93
Db 141 GATGAAGTAAACAGATCATGTGAACCAACCTGTGCTGAAGCAAAATCTGGAATGACTAC 200
QY 94 AsnLeuArgTyrAsnThrSerAspPheGlyGlyValIysAspLeuArgValProHis 113
Db 201 AAGCTGAATGGAAACCTCTGACTACCAAGGGGTGAGTTTCATGCGAGTCCCTGCAGAG 260
QY 114 ArgLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspSerThr 133
Db 261 AAGATCTGAAACACAGACATCGTCTTACCAACGCGCATGGGATTTCCAAGTGGAT 320
QY 134 TyrProThrAsnValValArgAsnAsnGlySerCysLeuTyrValProProGlyIle 153
Db 321 GACAAACCAAAAGCTCTACTCAAGTACACAGGAGAAAGTGAATTCCTCCGCGCATC 380
QY 154 PheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGlu 173
Db 381 TTTAAGAGCTCATGCAAAATCGAATGACCTACTTCCGTTTGACTACCAAACTGCACC 440
QY 174 MetLysPheGlySerTrpThrTyrAspGlyTyrGlnLeuAspLeuGlnLeuAspGlu 193
Db 441 ATGAAGTTGGCTCTCTGCTCTACGACAAAGGCAAGATCGACCTGGTCTCTCAITGGC-- 497
QY 194 GlyGlyAspLysSerSerPheValThrAsnGlyGluTrpGluLeuIleGlyValPro 213
Db 498 TCCTCCATGAACCTCAAGGACTATTGGGAAAGTGGCGAGTGGGCCATCATTAAGCCCG 557
QY 214 GlyLyArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPhe 233
Db 558 GGCCTAACATCAATCAAGTACACTGCTGTGAGGAGATCTACCAAGACATCAGTAC 617
QY 234 AlaValIleArgArgLysThrLeuTyrTyrPhePheAsnLeuIleValProCysVal 253
Db 618 TCGCTATACATTCCGCCCTCGCGCTGTCTACACCATCAACCTCATCTCCGTGCGCTG 677
QY 254 LeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeu 273
Db 678 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 737
QY 274 SerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThr 293
Db 738 ACCTCTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 797
QY 294 MetProAlaThrSerAspAlaValProLeuLeuGlyThrTyrPheAsnCysIleMetPhe 313
Db 798 ATCCCTTCCACCTCACTGGTCACTCCCTTGATCGGGGAGTACCTCTCTCTCTCTCTCT 857

```



```
QY 312 MetPheMetValAlaSerSerValValSerThrIleLeuIleLeuAsnTyrHisHisArg 331
Db 852 ATGTGTGTGTGCACCTTTTCATCGTCCACCACTGTGTGTCTCTCAATGTGCACACCGT 911
QY 332 HisAlaAspThrHisGluMetSerAspTyrIleArgCysValPheLeuTyrTrpLeuPro 351
Db 912 TCACCCAGCACTCACACATGGCATCTGGGTCAAGAGTGTCTTCCTGCACAAGCTGCC 971
QY 352 TrpValLeuArgMetSerArgProGlySerAlaThrThrProProAlaAlaArgValPro 371
Db 972 ACCTTCCTCTTCATGAAGCGCCTGGTGTGGAAGTCCAGC-----CCGCGCAGGTCCT 1025
QY 372 ProProProAspLeuGluLeu-----ArgGluArgSerSerLysSerLeuLeu----- 387
Db 1026 CATTCACGCCAGCTGCACCTTGACACAGCTGAAGCTACGTCCACCTCTGCGCTTAGGCC 1085
QY 388 -----AlaAsnValLeuAspIleAspAspPheArgHisProGlnAlaGlnGln 404
Db 1086 AGCAGCCCATCCACCTCTATGGGAATTCATGTACTTTGTGAACCTGTCCCTGCCACT 1145
QY 405 ProGln---CysCysArgTyrTyrArgGlyGlyGluGluAsnGlyAlaGlyLeuAlaAla 423
Db 1146 CCTAAGTCTGCAGTCAGCTCCACACGCGCAGGCTCCCGAGGATGCCCGGCTGAGGTCC 1205
QY 424 HisSerCysPheGlyValAspTyrGluLeuSerLeuLeuLeuGlyLeuValle 443
Db 1206 TCTGGGAGGTTCCGGCAAGAT-----CTACAGGAAGCATTAGAGGCGTCAGCTTCATC 1259
QY 444 ThrAspGlnMetArgLysAspAspGluAspAlaAspIleSerArgAspTyrLysPheAla 463
Db 1260 GCACAGCATTTGGAGATGATCGAGATCAGAGTCAGAGTCTATCGAGGACTGGAAATTCGTC 1319
QY 464 AlaMetValValAspArgLeuCysLeuIleIlePheThrLeuPheThrIleAlaThr 483
Db 1320 GCAATGTGTCTGCACCGCTGTCTCTGGGTGTCTGTATGTGTGTATCTGGGCACC 1379
QY 484 LeuAlaValLeuLeu 488
Db 1380 ATGGGGCTCTTCCTG 1394

RESULT 12
LOCUS AY402876
DEFINITION Homo sapiens CHRN4 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY402876
VERSION AY402876.1 GI:39758859
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1442)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL
PUBMED 14671302
REFERENCE 2 (bases 1 to 1442)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
```

```
source 1..1442
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1442
/gene="CHRN4"
/locus_tag="HCM1370"

ORIGIN
Alignment Scores:
Pred. No.: 6,24e-72 Length: 1442
Score: 895.50 Matches: 184
Percent Similarity: 58.60% Conservative: 92
Best Local Similarity: 39.07% Mismatches: 163
Query Match: 33.60% Indels: 32
DB: 29 Gaps: 7

US-09-303-232-4 (1-496) x AY402876 (1-1442)
QY 34 GluLysArgLeuLeuHisHisLeuLeuAsp-----HisTyrAsnValLeuGluArgPro 51
Db 24 GAGGAAAAGCTGATGGACGACCTTCTGACAAAAACCGGTTACATAAACCTGATCGGCCA 83
QY 52 ValValAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIle 71
Db 84 GCCACACAGCTCTCACAGCTCATCTCCATCAAGCTGCAGCTCTCCCTGGCCGACCTATC 143
QY 72 AspValAspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeuLysLeuGluTrpAsn 91
Db 144 AGCGTGAATGAGCGAGCAGATCATGACCACCAATGCTCTGGCTGAAACAGAGATGGACT 203
QY 92 AspMetAsnLeuArgTrpAsnThrSerAspPheGlyGlyValLysAspLeuArgValPro 111
Db 204 GATTACCGCGCTGATCTGGAACAGCTCCCGCTACGAGGCTGTGAACATCTCTGAGATCCCT 263
QY 112 ProHisArgLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp 131
Db 264 GCAAAGCGCATCTGGTTCCTGACATCGCTTTTACAAACGCCGACGGACCTATGAG 323
QY 132 SerThrTyrProThrAsnValValArgAsnAsnGlySerCysLeuTyrValProPro 151
Db 324 GTGCTCTCTACACCACTTGTAGTCCGCTCCAAACGCGACGCTCCTGTGGTCCCCCT 383
QY 152 GlyIlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArg 171
Db 384 GCCATCTACAAGAGCGCTCAAGATTGAGTGAAGTACTTCTCCCTTCGACCGACAGAAC 443
QY 172 CysGluMetLysPheGlySerTrpThrTyrAspGlyTyrGlnLeuAspLeuGln 191
Db 444 TGCACCTCAAGTTCCGCTCTCGACCTATGACACACGAGGATGAGACATGGTCTCATG 503
QY 192 AspGluGlyGlyAspIleSerSerPheValThrAsnGlyGluTrpGluLeuIleGly 211
Db 504 ACGCCC--ACAGCCAGCATGGATGACTTTACTCCCGAGTGTGTGAGTGGGACATATGGGCC 560
QY 212 ValProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTrpIleAspIle 231
Db 561 CTCCAGGGAGAGAGGACA-----GTGAACCCACAGACCCCGACCTACGTGGACGTG 611
QY 232 ThrPheAlaValValIleArgArgLysThrLeuTyrPhePheAsnLeuIleValPro 251
Db 612 ACTTACGACTTCATCATCAAGCGCAGGCTCTGTCTTACACCATCAACCTCATCATCCCC 671
QY 252 CysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProAspSerGlyGlu 271
Db 672 TGGTGTCTACCACTTGTGGCCATCTCTTCTTACTCGCATCCGACTGGGGGAG 731
QY 272 LysLeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAla 291
Db 732 AAGATGACATGTGCATCTCAGTGTCTGCTGGCACTGACATTCCTCTCTGCTCATCTCC 791
QY 292 GluThrMetProAlaThrSerAspAlaValProLeuLeuGlyThrTyrPheAsnCysIle 311
Db 311
```


Db 1291 CTGGAGTCTTTATGCTGGTGTCTCATCGGACGCTGGCTGTG 1335

Search completed: May 9, 2004, 04:20:29
Job time : 3493.3 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 12:59:49 ; Search time 5953.59 Seconds
(without alignments)
10942.062 Million cell updates/sec

Title: US-09-303-232-5_COPY_95_1597
Perfect score: 1503
Sequence: 1 atggccctatgttggcggc.....caccgcataatcatcgtgcaa 1503

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl.*
- 1: gb.ba.*
 - 2: gb.htg.*
 - 3: gb.in.*
 - 4: gb.om.*
 - 5: gb.ov.*
 - 6: gb.pat.*
 - 7: gb.ph.*
 - 8: gb.pl.*
 - 9: gb.pr.*
 - 10: gb.ro.*
 - 11: gb.sts.*
 - 12: gb.sy.*
 - 13: gb.un.*
 - 14: gb.vi.*
 - 15: em.ba.*
 - 16: em.fun.*
 - 17: em.hum.*
 - 18: em.in.*
 - 19: em.mu.*
 - 20: em.om.*
 - 21: em.or.*
 - 22: em.ov.*
 - 23: em.pat.*
 - 24: em.ph.*
 - 25: em.pl.*
 - 26: em.ro.*
 - 27: em.sts.*
 - 28: em.un.*
 - 29: em.vi.*
 - 30: em.htg.hum.*
 - 31: em.htg.inv.*
 - 32: em.htg.other.*
 - 33: em.htg.mus.*
 - 34: em.htg.pin.*
 - 35: em.htg.rod.*
 - 36: em.htg.mam.*
 - 37: em.htg.vrt.*
 - 38: em.sy.*
 - 39: em.htgo.hum.*
 - 40: em.htgo.mus.*
 - 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1503	100.0	3029	3	AF143847	AF143847 Heliothis
2	1503	100.0	3109	6	E58348	E58348 Nucleic aci
3	1503	100.0	3109	6	AX009614	AX009614 Sequence
4	770.2	51.2	2023	3	AF321447	AF321447 Drosophil
5	751	50.0	2023	3	AF321445	AF321445 Drosophil
6	751	50.0	2023	3	AF321446	AF321446 Drosophil
7	747.8	49.8	1865	3	DME554209	AJ554209 Drosophil
8	747.8	49.8	1699	3	BT011147	BT011147 Drosophil
9	715.2	47.6	2068	3	AF321448	AF321448 Drosophil
10	654	43.5	2110	3	AF321449	AF321449 Drosophil
11	575	38.3	3629	3	AF143846	AF143846 Heliothis
12	575	38.3	3700	6	AX009612	AX009612 Sequence
13	575	38.3	3701	6	E58347	E58347 Nucleic aci
14	535	35.6	1683	3	DME554210	AJ554210 Drosophil
15	521.6	34.7	2307	3	AF272778	AF272778 Drosophil
16	412.4	27.4	2834	3	AY036613	AY036613 Drosophil
17	411.4	27.4	1668	5	AY298752	AY298752 Takifugu
18	411.4	27.4	2886	6	E58346	E58346 Nucleic aci
19	411.4	27.4	2886	6	AX009610	AX009610 Sequence
20	378.8	25.2	2107	9	AF486623	AF486623 Macaca mu
21	377.2	25.1	1509	9	HSNACHRA7	Y08420 H.sapiens m
22	377.2	25.1	1916	9	AF385585	AF385585 Homo sapi
23	377.2	25.1	1977	9	HSU40583	U40583 Human alpha
24	376.8	25.1	1551	4	BTAT7	X93604 B.taurus mr
25	376.4	25.0	2106	10	RATNAPAD	L31619 Rattus ratt
26	376.4	25.0	3030	10	HS3987	X53987 nicotinic r
27	375.6	25.0	2087	9	HSARA7A	X70297 H.sapiens m
28	375	25.0	1559	9	HUMA7NAR	L25827 Human a7 ni
29	374	24.9	1509	6	AX054567	AX054567 Sequence
30	374	24.9	1876	6	AR282833	AR282833 Sequence
31	374	24.9	1876	6	AX719088	AX719088 Sequence
32	374	24.9	1876	9	HSU62436	U62436 Human nicot
33	374	24.9	2088	10	AF225980	AF225980 Mus muscu
34	373.6	24.9	1590	6	AR261850	AR261850 Sequence
35	373.6	24.9	1590	6	BD023656	BD023656 Variant h
36	373.6	24.9	1876	6	AR055255	AR055255 Sequence
37	373.6	24.9	1876	6	AR071403	AR071403 Sequence
38	373.6	24.9	1876	6	AR173187	AR173187 Sequence
39	373.6	24.9	1876	6	AR224030	AR224030 Sequence
40	372.4	24.8	1509	6	AX054575	AX054575 Sequence
41	372.4	24.8	1509	6	AX054577	AX054577 Sequence
42	372.4	24.8	1848	10	MUSNAES	L37663 Mus musculu
43	371.8	24.7	1665	5	AY298753	AY298753 Takifugu
44	370.8	24.7	1509	6	AX054579	AX054579 Sequence
45	365.8	24.3	1542	5	AY299466	AY299466 Takifugu

ALIGNMENTS

RESULT 1
AF143847
LOCUS AF143847
DEFINITION Heliothis virescens putative nicotinic acetylcholine receptor alpha
7-2 subunit mRNA, complete cds.
ACCESSION AF143847
VERSION AF143847.1 GI:4895006
KEYWORDS
SOURCE Heliothis virescens (tobacco budworm)
ORGANISM Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE 1 (bases 1 to 3029)
AUTHORS Schulte,T., Oellers,N. and Adamczewski,M.

TITLE Putative alpha subunits of insect nicotinic acetylcholine receptors more similar to vertebrate alpha 7 subunits and C. elegans Ce21 than to other insect nicotinic acetylcholine receptor alpha subunits

JOURNAL

REFERENCE

2 (bases 1 to 3029)

AUTHORS Schulte, T., Oellers, N. and Adamczewski, M.

TITLE Direct Submission

JOURNAL Submitted (19-APR-1999) ZP-BTB, Bayer AG, Bldg. Q 18,, Leverkusen

51368, Germany

FEATURES

source

1..3029

/organism="Heliothis virescens"

/mol_type="mRNA"

/db_xref="taxon:7102"

95..1600

/note="hvnachra7-2"

/codon_start=1

/product="putative nicotinic acetylcholine receptor alpha

7-2 subunit"

/protein_id="AAD32698.1"

/db_xref="GI:4895007"

/translation="MADMLAALALLPVSQEPHEKRLNALLANYNLTLPVANE
SELEVFGLTQIIDIIVDEKNQLLIINWLEWYNDYLRWDSYGGVKDLRITPN
KLMPDVLMTNSADEPDGYQTNVVRSGSCLYVPPGIFKSKCKMDIAWFPDDH
CDMKFGSWTDGNDLVDLDEAGDLSDFITNGEWLYLIGMPGKNTITVACCPYV
DVTFTIMRRRTLYFENLIVPCVLISSMALLGFTLPDSEKLTGLVTLLSITVEL
NLVAETIPQVSDAIPLLGTFFNCIMFNVASSVLTVVLYNTHRTADIHMPQWIKSV
FLOWPLWLNRSRPGKTIKRTIMNTRMELEKRSKSLANLVLDIDDDPRHGP
PFNSTAGLWPGSCSTFRDPRFSVPRSTMEDVGGGLSHRELHLILRELQFITA
RMKKADEABLIIDWKPMVAVDRFCLFVETLETTIATVAVLISAPHIIQV"

ORIGIN

Query Match 100.0%; Score 1503; DB 3; Length 3029;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1503; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTATGTTGGCGGCTTGGCGCTGTGCTGCTTGTGCTGCCGCTATCGGACCAAGGT 60
DB 95 ATGGCCCTATGTTGGCGGCTTGGCGCTGTGCTGCTGCTTGTGCTGCCGCTATCGGACCAAGGT 154

QY 61 CCTCAGAGAGAGACTCTGAAACGCTTGTGCGAACTACAAACCCCTGGAGCGACCG 120
DB 155 CCTCAGAGAGAGACTCTGAAACGCTTGTGCGAACTACAAACCCCTGGAGCGACCG 214

QY 121 GTGCCAACAGAGCGAACCGCTAGAGTTCAGGTTGCGCTTGACCTTGCAGCAATCAT 180
DB 215 GTGCCAACAGAGCGAACCGCTAGAGTTCAGGTTGCGCTTGACCTTGCAGCAATCAT 274

QY 181 GAGTGGACGAGAGAACTCAACTTATTAACCAATATATGGCTGTGCTTGGAGTGAAT 240
DB 275 GAGTGGACGAGAGAACTCAACTTATTAACCAATATATGGCTGTGCTTGGAGTGAAT 334

QY 241 GACTACAACTGAGGTGGAACGACGAGATATGGCGGGTCAAGACCTCAGATCAGC 300
DB 335 GACTACAACTGAGGTGGAACGACGAGATATGGCGGGTCAAGACCTCAGATCAGC 394

QY 301 CCCAACAGTTGTGGAAGCGGAGCTCTTATGTAATATAGTGTGACGAGGGTTTGAC 360
DB 395 CCCAACAGTTGTGGAAGCGGAGCTCTTATGTAATATAGTGTGACGAGGGTTTGAC 454

QY 361 GGGACCTTACAGACCAACGTTGGTTCAGAAAGCGGCGGAGTTGCCCTGTACGTGCCACCT 420
DB 455 GGGACCTTACAGACCAACGTTGGTTCAGAAAGCGGCGGAGTTGCCCTGTACGTGCCACCT 514

QY 421 GGCATATTCAGAGCAGATGAGATGAGATGCGGTGGTTCCCTTCGACGACCAACAC 480
DB 515 GGCATATTCAGAGCAGATGAGATGAGATGCGGTGGTTCCCTTCGACGACCAACAC 574

QY 481 TGTGATATGAGTTCCGTTAGCTGGACATATGACGGCAATCAGTTGGATCTGGTCTAAAA 540
DB 575 TGTGATATGAGTTCCGTTAGCTGGACATATGACGGCAATCAGTTGGATCTGGTCTAAAA 634

QY 541 GATGAGCAGCGCGGATCTATCGACTTCAATAACAATGGGAGTGTATCTAATAGGA 600
DB 635 GATGAGCAGCGCGGATCTATCGACTTCAATAACAATGGGAGTGTATCTAATAGGA 694

QY 601 ATGCCAGGCAAAAAGAACACAATAACATACGCTGCTGCCCGAGCCCTACGTGACCTC 660
DB 695 ATGCCAGGCAAAAAGAACACAATAACATACGCTGCTGCCCGAGCCCTACGTGACCTC 754

QY 661 ACCTTCACCATCATGATAAGAGACGAACTTGTACTTCTTCAACCTCATCGTCCG 720
DB 755 ACCTTCACCATCATGATAAGAGACGAACTTGTACTTCTTCAACCTCATCGTCCG 814

QY 721 TGGCTGTGATCTCATGATGGCACTCCTCGGCTTACACTGCCACAGACTCCGGAGAG 780
DB 815 TGGCTGTGATCTCATGATGGCACTCCTCGGCTTACACTGCCACAGACTCCGGAGAG 874

QY 781 AAACCTCACACTTGGAGTCACTTATCTTCTATCGCTGACGGTGTCTCTCAACCTGATGCC 840
DB 875 AAACCTCACACTTGGAGTCACTTATCTTCTATCGCTGACGGTGTCTCTCAACCTGATGCC 934

QY 841 GAGACCTGCCACAGGCTCTCGGAGCTATCCCTCTGTAGGACGTACTTCAATGATC 900
DB 935 GAGACCTGCCACAGGCTCTCGGAGCTATCCCTCTGTAGGACGTACTTCAATGATC 994

QY 901 ATGTTCACTGTAGCGTCTGTGCTACTGACTGTGTGTGTACTCAATTACCAACATCGA 960
DB 995 ATGTTCACTGTAGCGTCTGTGCTACTGACTGTGTGTGTACTCAATTACCAACATCGA 1054

QY 961 ACAGCTGATATACATGAAATGCCACAGTGGATAAAATCAGTATTCTTCAATGTTGCCA 1020
DB 1055 ACAGCTGATATACATGAAATGCCACAGTGGATAAAATCAGTATTCTTCAATGTTGCCA 1114

QY 1021 TGGATCTCGAATGTGCGAGCCGAGGAAGAGATCACAGGAGACTATATGATGATGAC 1080
DB 1115 TGGATCTCGAATGTGCGAGCCGAGGAAGAGATCACAGGAGACTATATGATGATGAC 1174

QY 1081 ACAGGATGAGGAGCTTGGAACTGAAGAGAGGTCGCGAAGTCTTGTGGCGCAATGTT 1140
DB 1175 ACAGGATGAGGAGCTTGGAACTGAAGAGAGGTCGCGAAGTCTTGTGGCGCAATGTT 1234

QY 1141 CTAGATATGATGATGATCTTTCAGACACGCGCCCTCCGCTCTCTAACAGTACTGCTCGAC 1200
DB 1235 CTAGATATGATGATGATCTTTCAGACACGCGCCCTCCGCTCTCTAACAGTACTGCTCGAC 1294

QY 1201 GGGAAATTTGGACCTGGGTGCTCAATATTCGACACGATTTCCGTGGTGGTTCGTCGT 1260
DB 1295 GGGAAATTTGGACCTGGGTGCTCAATATTCGACACGATTTCCGTGGTGGTTCGTCGT 1354

QY 1261 CCGTCCACGATGAAAGACGTCGGCGGCTGGGTAGCCACCATCCGAGCTCACCTC 1320
DB 1355 CCGTCCACGATGAAAGACGTCGGCGGCTGGGTAGCCACCATCCGAGCTCACCTC 1414

QY 1321 ATACTGAGAGCTGCAGTTTCATCAGCGCCAGGATGAAGAGGCTGATGAGGAAGCCGAG 1380
DB 1415 ATACTGAGAGCTGCAGTTTCATCAGCGCCAGGATGAAGAGGCTGATGAGGAAGCCGAG 1474

QY 1381 CTGATCAGCAGCTGGAAGTTTGTGCGATGTTGTTGATAGGTTTTCCTGTTGCTGTTTC 1440
DB 1475 CTGATCAGCAGCTGGAAGTTTGTGCGATGTTTGTGATAGGTTTTCCTGTTGCTGTTTC 1534

QY 1441 ACACTTTTCAATCATCGGACAGTACGTCTCTGTTATCGGCACCGCATATCATCGTG 1500
DB 1535 ACACTTTTCAATCATCGGACAGTACGTCTCTGTTATCGGCACCGCATATCATCGTG 1594

QY 1501 CAA 1503
DB 1595 CAA 1597

RESULT 2

E58348

LOCUS

DEFINITION Nucleic acid encoding insect actyl choline receptor subunit.

E58348 3109 bp DNA linear PAT 18-JUN-2001

Accession	E58348
Version	E58348.1
Keywords	GI:13019347
Source	JP 2000023680-A/3.
Organism	Heliothis virescens (tobacco budworm)
	Heliothis virescens
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea; Noctuidae; Heliothinae; Heliothis.
Reference	1 (bases 1 to 3109)
Authors	Martin,A., Nadjai,E. and Thomas,S.
Title	Nucleic acid encoding insect acyl choline receptor subunit
Journal	Patent: JP 2000023680-A 3 25-JAN-2000;
	BAYER AG
Comment	OS Heliothis virescens PN JP 2000023680-A/3 PD 25-JAN-2000 PF 26-APR-1999 JP 1999118159 PR 04-MAY-1998 DE 19819829.9 PI MARTIN ADAMTOSUEUSUKI,NADJAI ERASU,THOMAS SCHULTE PC C12N15/09,A01K67/033,C07K16/28,C12N1/21,C12N5/10,PC C12Q1/68, PC G01N33/15,G01N33/50//(C12N1/21,C12R1:19),C12N15/00,C12N5/00 CC
Features	FT Key Location/Qualifiers
source	Location/Qualifiers 1..3109 /organism="Heliothis virescens" /mol_type="genomic DNA" /db_xref="taxon:7102"
Origin	
Query Match	100.0%; Score 1503; DB 6; Length 3109;
Best Local Similarity	100.0%; Pred.No. 0;
Matches 1503; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ATGCGCCCTATGTGGCGCCCTTCGGCGTGCTGCCTTTGCTGCCGTATCGAGCAAGGT 60
Db	95 ATGCGCCCTATGTGGCGCCCTTCGGCGTGCTGCCTTTGCTGCCGTATCGAGCAAGGT 154
Qy	61 CCTCACGAGAGACTCTCTGAACGCGTTGCTGGGAACTACACACCCTGGAGCGACG 120
Db	155 CCTCACGAGAAGACTCTCTGAACGCGTTGCTGGGAACTACACACCCTGGAGCGACG 214
Qy	121 GTGGCCCAACGAGAGCAACCGCTTAGAGGTCAAGGTCGGCTTGA CTTGCGACCAATCAT 180
Db	215 GTGGCCCAACGAGAGCAACCGCTTAGAGGTCAAGGTCGGCTTGA CTTGCGACCAATCAT 274
Qy	181 GACGTGACGAGAAGATCAACTATTATAACCAATATATCGCTGTG TCGGTGGAAT 240
Db	275 GACGTGACGAGAAGATCAACTATTATAACCAATATATCGCTGTG TCGGTGGAAT 334
Qy	241 GACTACAACCTGAGGTGGAACGACGAGAGTATGGCGGGTCAAGG ACTCAGATCACG 300
Db	335 GACTACAACCTGAGGTGGAACGACGAGAGTATGGCGGGTCAAGG ACTCAGATCACG 394
Qy	301 CCCAACAGTTGTGGAACCGGACGCTCTTATGTATATATAGTGTG ACAGGGTTTTGAC 360
Db	395 CCCAACAGTTGTGGAACCGGACGCTCTTATGTATATATAGTGTG ACAGGGTTTTGAC 454
Qy	361 GGGACCTACGACCAACGCTGGTGGTCAAGACGGCGGCGGAGTGC TTGACGTGCCA 420
Db	455 GGGACCTACGACCAACGCTGGTGGTCAAGACGGCGGCGGAGTGC TTGACGTGCCA 514
Qy	421 GGCATATTCAGAGACATGCAAGATGACATCGCGTGGTTCCCTTC GAGCGACCAAC 480
Db	515 GGCATATTCAGAGACATGCAAGATGACATCGCGTGGTTCCCTTC GAGCGACCAAC 574
Qy	481 TGTGATATGAAGTTCGGTAGCTGGACATATGACGGCAATCAGTTG GATCTGGTGCTAAA 540
Db	575 TGTGATATGAAGTTCGGTAGCTGGACATATGACGGCAATCAGTTG GATCTGGTGCTAAA 634
Qy	541 GATGAGGCGGCGGCGATCTATCGGACTTCAATACAAATGGGAGT GGTATCTAATAGGA 600

AX009614.1 GI:9996846
Heliothis virescens (tobacco budworm)
Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
1
Adamczewski, M.D., Schulte, T.D. and Oellers, N.D.
Nucleic acids encoding acetylcholin-receptor subunits from insects
Patent: EP 0962528-A 5 08-DEC-1999;
BAYER AG (DE)
FEATURES
source
1. .3109
/organism="Heliothis virescens"
/mol_type="unassigned DNA"
/db_xref="taxon:7102"
95. .1600
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07501.1"
/db_xref="GI:9996847"
/db_xref="REMTREMBL:CAC07501"
/translation="MAPMLAALALLLPVSBQPHKRLNALLNANTLERPVANE
SEPLEVRFGLTQIIDVDEKNQLLIINLSLEWINDYNLRWDSYGVKDLRIITPN
KLWPKDVLMTNSADGPDGYQTNVVRSGSLVPPGIFKSKCKMDIAWFPDDOH
CDMKFGSYTDNQLDLVXDEAGDLSDRITNGEWFLIGMPGKNITITVACCPYV
DVFTIMRERTLYAFENLJPCVLISMSALLGFTLPDPSGEXLTLGVILLSLTVFL
NLVATILPQVSDAIPLLGTFCNFMFVASSVLTVVLYNHHRTADIHMPQWIKSV
FLOWPLMIRSPGKIKTRKIMMTRNRELEKRSKSLLANLVIDDDFRHGP
PPNASTGNLPGCSIFTRDFRSPVPMEDVGGGLSHHREHLILRELFQITA
RMKKADEABELISDWKFAAMVVDVDFCLFVETLFTIATVAVLISAPHIIQV"
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 1503; DB 6; Length 3109;
Matches 1503; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCCCTATGTTGGCGCCTTGGCGCTGTGCTGCTTGTGCTGCCGCTATCGAGCAAGGT 60
DB 95 ATGGCCCTATGTTGGCGCCTTGGCGCTGTGCTGCTTGTGCTGCCGCTATCGAGCAAGGT 154
QY 61 CCTCAGAGAGAGACTCTCTGAACGCTTGTGCGAACTACAAACCTTGGAGCGACCG 120
DB 155 CCTCAGAGAGAGACTCTCTGAACGCTTGTGCGAACTACAAACCTTGGAGCGACCG 214
QY 121 GTGCGCAACAGAGCGAACCGCTAGAGGTGAGTTGCGCTTGCCTTGCAGCAAAATCAT 180
DB 215 GTGCGCAACAGAGCGAACCGCTAGAGGTGAGTTGCGCTTGCAGCAAAATCAT 274
QY 181 GAGTGGACGAGAGAACTCACTACTTATAACCAATATATGGCTGTGCTTGGAGTGAAT 240
DB 275 GAGTGGACGAGAGAACTCACTACTTATAACCAATATATGGCTGTGCTTGGAGTGAAT 334
QY 241 GACTCAACCTGAGTGGAGACGACGAGATATGGCGGGTCAAGGACCTCAGGATCACG 300
DB 335 GACTCAACCTGAGTGGAGACGACGAGATATGGCGGGTCAAGGACCTCAGGATCACG 394
QY 301 CCCAAACAAGTTGTGGAAGCGGACGCTCTTATGTAATAGTGTGACGAGGGTTTGAC 360
DB 395 CCCAAACAAGTTGTGGAAGCGGACGCTCTTATGTAATAGTGTGACGAGGGTTTGAC 454
QY 361 GGGACCTTACAGACCAACGTTGGTGTGAGAGCGGGCGGAGTTGCCCTGTACGTGCCACCT 420
DB 455 GGGACCTTACAGACCAACGTTGGTGTGAGAGCGGGCGGAGTTGCCCTGTACGTGCCACCT 514
QY 421 GGCATATTCAAGAGCACATCGAGATGGACATCGCGTGGTTCCCTTCGACGACCAACAC 480
DB 515 GGCATATTCAAGAGCACATCGAGATGGACATCGCGTGGTTCCCTTCGACGACCAACAC 574
QY 481 TGTGATATGAAGTTTCGGTAGCTGGACATATGACGGCAATCAGTTGGATCTGGTCTAAAA 540
DB 575 TGTGATATGAAGTTTCGGTAGCTGGACATATGACGGCAATCAGTTGGATCTGGTCTAAAA 634

QY 541 GATGAGCAGCGCGGATCTATCGGACTTCTATACAAATCGGAGTGGTATCTAATAGA 600
DB 635 GATGAGCAGCGCGGATCTATCGGACTTCTATACAAATCGGAGTGGTATCTAATAGA 694
QY 601 ATGCCAGCAAAAAGAACACAATAACATACGCGTGTGCTGCCCGAGCCCTAGCTGACGTC 660
DB 695 ATGCCAGCAAAAAGAACACAATAACATACGCGTGTGCTGCCCGAGCCCTAGCTGACGTC 754
QY 661 ACCTTCACCATCATGATAAGAACGAACTTGTGTACTTCTTCAACTGATCGTCCG 720
DB 755 ACCTTCACCATCATGATAAGAACGAACTTGTGTACTTCTTCAACTGATCGTCCG 814
QY 721 TGGTGTGTGATCTCATCGATGGCACTCTCGGCTTCACTGCGCACAGATCCGGAGAG 780
DB 815 TGGTGTGTGATCTCATCGATGGCACTCTCGGCTTCACTGCGCACAGATCCGGAGAG 874
QY 781 AAACCTCACACTTGGAGTCACTATTCTTCTATCGCTGACGGTGTTCCTCAACTGGTAGCC 840
DB 875 AAACCTCACACTTGGAGTCACTATTCTTCTATCGCTGACGGTGTTCCTCAACTGGTAGCC 934
QY 841 GAGACCTGCCACAGGTCCTCGAGGCTATCCCTCTGTAGGAGCGTACTTCAATTGCAATC 900
DB 935 GAGACCTGCCACAGGTCCTCGAGGCTATCCCTCTGTAGGAGCGTACTTCAATTGCAATC 994
QY 901 ATGTTCAATGTTAGCGTCTGTGCTACTGACTGCTGCTGCTGCTCAATTACCACATCGA 960
DB 995 ATGTTCAATGTTAGCGTCTGTGCTACTGACTGCTGCTGCTCAATTACCACATCGA 1054
QY 961 ACAGCTGATATACATGAAATGCCACAGTGGATAAAATCAGTATTCTTCAATGGTTGCCA 1020
DB 1055 ACAGCTGATATACATGAAATGCCACAGTGGATAAAATCAGTATTCTTCAATGGTTGCCA 1114
QY 1021 TGGATCTGCAATGTGAGCGCCAGGAAAGATACACGAGGAAGACTATAATGATGAAC 1080
DB 1115 TGGATCTGCAATGTGAGCGCCAGGAAAGATACACGAGGAAGACTATAATGATGAAC 1174
QY 1081 ACAGGATGAGGAGCTGGAACCTGAAGGAGAGTCTGCGAAGTCTTGTGCTGGCGAATGTT 1140
DB 1175 ACAGGATGAGGAGCTGGAACCTGAAGGAGAGTCTGCGAAGTCTTGTGCTGGCGAATGTT 1234
QY 1141 CTAGATATTGATGATGACTTTCAGACACGCGCCCTCCGCTCTCTAACAGTACTGCTCGACC 1200
DB 1235 CTAGATATTGATGATGACTTTCAGACACGCGCCCTCCGCTCTCTAACAGTACTGCTCGACC 1294
QY 1201 GGGAAATTTGGGACCTGGGTGCTCAATATTCCGACCGAATTTCCGTCGCTGCTCGCT 1260
DB 1295 GGGAAATTTGGGACCTGGGTGCTCAATATTCCGACCGAATTTCCGTCGCTGCTCGCT 1354
QY 1261 CCGTCCACGATGGAAGACGTGGCGGGCTGGGTAGCCACCATCGCGAGCTGCACCTC 1320
DB 1355 CCGTCCACGATGGAAGACGTGGCGGGCTGGGTAGCCACCATCGCGAGCTGCACCTC 1414
QY 1321 ATACTGAGAGAGCTGCAGTTTCATCAGCGCCAGGATGAAGAGGCTGATAGGAGCCGAG 1380
DB 1415 ATACTGAGAGAGCTGCAGTTTCATCAGCGCCAGGATGAAGAGGCTGATAGGAGCCGAG 1474
QY 1381 CTGATCAGGACTGGAAGTTTGTGCGATGGTTGTTGATAGGTTTGGCTGTTGCTGTTTC 1440
DB 1475 CTGATCAGGACTGGAAGTTTGTGCGATGGTTGTTGATAGGTTTGGCTGTTGCTGTTTC 1534
QY 1441 ACACCTTTTCAACATCATCGACAGTAGCTGCTGTTATCGGCACCGCATATCATCGTG 1500
DB 1535 ACACCTTTTCAACATCATCGACAGTAGCTGCTGTTATCGGCACCGCATATCATCGTG 1594
QY 1501 CAA 1503
DB 1595 CAA 1597

RESULT 4
AF321447
LOCUS

2023 bp mRNA linear INV 29-APR-2002


```
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
1. .1665
/gene="nAcRalpha-30D"
44. .1528
/codon_start=1
/product="nicotinic acetylcholine receptor subunit
Daiphae"
/protein_id="CAD86935.1"
/db_xref="GI:29466435"
/translation="MDSPLASLSFLVLLIFLAIKESCOGPKHEKRLNHLISNYTL
ERPANESEPLEVKEFGLTLOQI IDVDEKNQLLINLWLSWYNDYNLRWNETYGVK
DLRITPNKLPVLMNTSADGPDGTHSVVVGSGSLYVPPGIFPKSTCKMDITW
FPDDQHCWKFSWTDGNQDLVLSSEDDGSDSDFITNGEYLLAMPKGNITIVA
CSPEYVDITFTQIRRTLYFFNLIVPCVLISSMALLGFTLPDPSGEKLITGLVTL
LSUTFLNLVAESPSTSDAVPLIGTVFNCIMFVASSVLTVVLNHYHRTADIDEM
PWTKSVELQWLPWILRMGRGRKI TRKTI LLSNRKLELKERSSKLLANVLIDID
DFBHTISGQTAIGSSASRGRPTVEEHHTAIGCNHKLHLILKELOFITARMKADD
EALIGDKWFAAMVVDRLFLVFLFIITAIIVTLLSAPHILVQ"
44. .121
/gene="nAcRalpha-30D"
122. .1525
/gene="nAcRalpha-30D"
/product="nicotinic acetylcholine receptor subunit
Daiphae"
Query Match 49.8%; Score 747.8; DB 3; Length 1665;
Best Local Similarity 69.9%; Pred. No. 1e-187;
Matches 1044; Conservative 0; Mismatches 407; Indels 42; Gaps 1;
QY 11 TGTGGCGGCTTGGCGCTGCTGCTTTGCTGCCGCTATCGAGCAAGGTCCTCACGAGA 70
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
75 TGTGTGCTGCTGATCTTTCTGGCGATAATTAAGAAAGAGCTGTCAAGGAGCCTCATGAAA 134
QY 71 AGAGACTCCTGAACGCGTCTGTCGGAATACACACCCCTGGAGCGCCGGTGGCCAAAG 130
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
135 AGCGCTGCTGAACCACTGCTGTCCACCTACAAATACGCTGGAGCGACCCGTGGCCAAATG 194
QY 131 AGAGCGAACCGCTAGAGCTCAGGTTCCGCTTGACCTTGACGCAAAATCAITGACGTGACG 190
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
195 AATCGAGCCCTCGAGGTTAAGTTCCGACTGACGCTGACGAGATCATCGAGTGGAGC 254
QY 191 AGAAGATCAACTTATTAACCAATATATGCTGTCTGTTGGAGTGAATGACTACAAAC 250
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
255 AGAAGATCAGCTTCTATACCAATCTTTGGCTTTCTGTTGGAGTGAACGACTACAAATC 314
QY 251 TGAGGTGAACGACGACGAGTATGGCGGGTCAAGGACCTCAGGATCAAGCCCAACAGAT 310
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
315 TGGCTGGAATGAACCGGAATACGCGGGGTCAAGGATCTACGAATCACGCCCAACAAAGC 374
QY 311 TGTGAAGCCGAGCGTCTTATGTATATAGTGTGACGAGGGTTTTGACGGGACCTAAC 370
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
375 TGTGAAGCCGAGCTCTCATGTACACAGCGGGATGAGGATTCGATGGCACGCTATC 434
QY 371 AGACCAACGTGTGTGACAGCGCGGAGTTCCTGTATCGTCCACCTGGCATATTCATCA 430
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
435 ACACGAGCGTGTGGTCAACATGCGCGGCGAGTTGTCTGTACGTGCCCCCTGGTATCTCA 494
QY 431 AGAGCACATCAAGATGGACATCGGTGTTTCCCTTCGACGACCAACACTGTGTATGA 490
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
495 AGAGCACATCAAGATGGACATACGTGTTTCCCTTCGATGACCAATTTGGGAAATGA 554
QY 491 AGTTGGTAGCTGGACATATGACGGCAATCAGTTGGATCTGGTGTCAAAAAGATGAGGACG 550
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
555 AATTCGAGTGTGAGCTTACGATGGAATCAGTTGGATTTGGTTTGGTTCCGAAGATG 614
QY 551 GCGCGAGTCTATCGACTTTCATACAAATGGGAGTGGTATCTAATAGGAATGCCAGCA 610
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
615 GAGGGGATCTTCCGATTTTCATAACAAATGCGAGTGGTACTTGTCTGGCATGCCGGAA 674
```

ORIGIN

```
611 AAAGAACAACAATAACATACGCGTCTGCCCGAGCCCTACGTGAGCGTCACTTACCA 670
675 AGAAGAAATAGATAGTCTACGCTGTGCCCGAGAACATATGTGATATACCTTTACTA 734
671 TCATGATAAGAAGACGAACCTTGTACTACTTCTTCAACCTGATCGTCCGCGCTGCTGA 730
735 TACAAATTCGTGCGCGTACATATATATTTTTCATTAATTTGTCATGTGTGCTTAA 794
731 TCTCATCGATGGCACTCTCTCGGCTTACACTGCGCACAGACTCCGAGAGAAAATCAAC 790
795 TCTCATCGATGGCCCTACTGGGCTTACATTTGCCGCGGATTTGGGCGAGAAAATGACGC 854
791 TTGGAGTCACTATTTCTTATCGTGTAGCGGTGTCTTCAACCTGGTGTAGCCGAGACCTGCG 850
855 TGGGCGTAACTATCTACTATCAATTAACAGTATTTCTTAAACCTTGTGCGGAGTCCATGC 914
851 CACAGGCTCTCGAGCGTATCCCTCTGTAGGAGCTACTTCAATTTGTCATGATGTTTATGG 910
915 CGACAACTCGGATGCTGTCTCTTATAGCACCTACTTCAATTTGTCATGATGTTTATGG 974
911 TAGCGTCTGTGGTACTGACTGTGGTGTACTCAATTTACCACCATCGAACAGCTGATA 970
975 TCGCCTCGTGGTGTGCTGACAGTAGTGTGTCTCAACTACCACCATCGACACGCGACA 1034
971 TACATGAATGCCACAGTGTGATTAATCAGTATTTCTTCAATTTGTCGATGATGCTGC 1030
1035 TTACAGAGATGCCACCGCTGGATCAAGTCCGTTTTTCTTCAATTTGCTCCCTGGATCTGC 1094
1031 GAATGTCTGAGCGCCAGGAAGATCACAGGAGACTAATAATGATGAACACGAGGATGA 1090
1095 GAATGGGTCCACCGGTCGCAAGATTACACGCAAAACAATCTATTAAAGCAATCGATGA 1154
1091 GGGAGCTGGAACCTGAAGAGAGGTGCTGGAAGTCTTGTGGCGAATGTTCTAGATATTG 1150
1155 AGGAGCTGGAGCTAAAGAGCGCTCTCTCCAAATCCCTGCTGGCCAATGTCTCGACATCG 1214
1151 ATGATGACTTCAGACACGCGCCCTCCGCTCTCTAACAGTACTGCTCGACCGGAAATTGG 1210
1215 ACAGCACTTCCGCGCACAAATATCTG----- 1241
1211 GACCTGGTCTCTCAATATTCCGACGAGATTTCGCTGGTCTGTCGTCGTCGTCACCA 1270
1242 -----GCTCCCAACCGCATTTGGCTGTCGGCCAGCTTGGTGGGCCCAACAACGG 1292
1271 TGAAGACGTGGCGCGGCTGGGTAGCCACCATCCGAGCTGCACTCATATCTGAGAG 1330
1293 TGGAGGAGCATCACACGGCCATCGGCTGCAATCACAAAGATCTTCATTAATTCTCAAAG 1352
1331 AGCTGAGTTCATCACGCGCCAGGATGAAGAGGCTGATAGGAGCGGAGCTGATCAGCG 1390
1353 AATTGCAATTTATTACGGCGCGGATCGCAAAAGCTGACGACGAAGCGGAATTGATCGCG 1412
1391 ACTGGAAGTTTGTGCGATGTTTGTATAGGTTTTGCTGTTTCCCTGTTTCCGATTTTCA 1450
1413 ATTGGAAGTTCGGCGCAATGTTTGTGGATAGATTTTGTAAATTTTTCACGCTCTTCA 1472
1451 CAATCATCGCAGCAGTAGTGTCTCTGTATPCGCAACCGCATATCATGTCGCAA 1503
1473 CGATTATTGCAACGGTTACGGTGTCTCTCCGCTCCGACATAAATCGTCAA 1525
```

RESULT 8
BT011147
LOCUS
DEFINITION
Drosophila melanogaster GH15518 full insert cDNA.
ACCESSION
BT011147.1
VERSION
BT011147.1
KEYWORDS
FLI CDNA.
SOURCE
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

830 TCAACAGTAGCGGAGTGTCTGTACGTGCCCTCGTGTATCTTCAAGAGCACATGCAAGA 889
446 TGGACATCGGTGTTCCCTTCAGACCAACACTGTGATGATGAAGTTCGGTAGCTGGA 505
890 TGGACATCAGCTGTCCCTTTCATGACCAACATGCGAATGAAATTCGGTAGTGGGA 949
506 CATATGACGGCAATCAGTTGATCTGTGCTTAAAGATGAGCGAGGCGGCATCTATCGG 565
950 CTTACGATGGAATCAGTTGATTTGGTTTGAATCCGAAGATGGAGGGATCTTCCG 1009
566 ACTTCATAACAAATGGGAGTGTATCTAATAGGAATGCCAGGCAAAAGAACACATAA 625
1010 ATTTTATAACAAATGGGAGTGTACTTCTGTCATGCCGGAAGAAAGATACGATAG 1069
626 CATACGGTGTGCTGCCGAGCCCTACGCGGAGCTCACCTTCACCATCATGATAGAAGAC 685
1070 TCTACGCTGTGCCGGAACCATATGTCGATATCACCTTTACTATATACAAATTCGTGCC 1129
686 GAACTTTGTACTTCTTCAACCTGATCGTCCGTCGTGCTGATCTCATCGATGGCAC 745
1130 GTACATATATATTTTTCATTTAATCGTGCCATGTGCTAATCTCATCGATGGCCC 1189
746 TCCTCGGCTTCACATGCGACAGACTCCGGAGAGAACTCACACTTGGAGTCACTATTC 805
1190 TACTGGGCTTCACATGCGCGGATTCGGCGAGAACTGACGCTGGGAGTTACAATTC 1249
806 TTCTATCGCTCAGCGTGTTCCTCAACCTGTAGCGGAGACCTGCGACAGTCTCCGACG 865
1250 TTCTATCGCTCAGAGTGTTCCTCAACCTGTAGCTGAGACATGCCCCAAGTACTGATG 1309
866 CTATCCCGCTGTAGGAGCTACTTCAATTCGATCATGTTTCATGCTAGCGTCTGCTGG 925
1310 CAATCCCGCTGTAGGACCTACTTCAATTCGATCATGTTTCATGCTGCTGCTGGTGG 1369
926 TACTGACTGTGGTGTACTCAATACCACTGAGCAAGCTGATATATCAATGAATGCCAC 985
1370 TGCTGACAGTGTGCTGCTCACTACCACTGAGCAAGCTGAGCAATTCACGAGATGCCAC 1429
986 AGTGATAAATCAGTATTCTTACAAATGTTGCTGCTGATGATGCTGCGAGCCAG 1045
1430 GTGGATCAAGTCCGTTTCTTCAATGCTGCTGCTGCTGATGCTTGGCAATGGTCCACCG 1489
1046 GGAAGAAGATCACAGGAAGCTATTAATGATGAACACAGGATGAGGAGCTGGAATGGA 1105
1490 GTCCGAAGATTACACGCAAAACAATACTATTAAGCAATCGCATGAAGGAGCTGGAGCTAA 1549
1106 AGGAGGTCGTCGAAGTCTTCTGCTGGCAATGTTCTAGATATTCATGATGACTTCAGAC 1165
1550 AGGAGGCTCTCTCCAAATCCCTGCTGGCAATGCTCTGACATCGACGAGCTTTCGGC 1609
1166 ACGGCCCTCCGCTCTTAAACAGTACTGCTCGACCGGAAATTTGGGACCTGGGTGCTCAA 1225
1610 ACACAAATCTG-----GCTCCC 1627
1226 TATTCGCGCAGGATTCGCTGGTTCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1285
1628 AAACCGCATTTGGCTCGTGGCCAGCTTGGTGGCCCAACACCGTGGAGGAGCATCA 1687
1286 GCGGCTGGTAGCACCATCGGAGCTGCATCTATGAGAGCTGCGAGTGCAGTTCATCA 1345
1688 CGGCCATCGGCTGCAATCAAAAGATCTTCACTAATTTTAAAGAAATTCGAATTTATTA 1747
1346 CCGCCAGGATGAAGAGCTGATGAGGAAGCGGAGCTGATCAGCGACTGGAAAGTTGCTG 1405
1748 CCGCGCGGATGCGCAAGCTGACGACGAGCGGAATTCATCGCGATTTGGAAGTTTCGCG 1807
1406 CGATGTTGTTGATAGGTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1465
1808 CAAATGTTGATAGATTTTGTGTTAAATGTTTTCACGCTCTTCACGATTTATGCAACGG 1867
1466 TAGCTGTCTGTTTATCGGACCGCATATCATCTGTCGAA 1503
1868 TTACGGTGTCTCTCCGCTCCGACATATCTGTCGAA 1905

RESULT 10
AF321449
LOCUS
DEFINITION
AF321449
2110 bp mRNA linear INV 29-APR-2002
Drosophila melanogaster nicotinic acetylcholine receptor Dalpha6
subunit variant type V (nAcRalpha-30D) mRNA, complete cds,
alternatively spliced.
ACCESSION
AF321449
VERSION
AF321449.1 GI:20152852
KEYWORDS
Drosophila melanogaster (fruit fly)
SOURCE
Drosophila melanogaster
ORGANISM
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 2110)
Grauso, M., Keenan, R.A., Culetto, E. and Sattelle, D.B.
Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
Dalphas, Dalphas6 and Dalphas7, in Drosophila melanogaster Identify a
New and Highly Conserved Target of Adenosine Deaminase Acting on
RNA-Mediated A-to-I Pre-mRNA Editing
Genetics 160 (4), 1519-1533 (2002)
21969411
MEDLINE
11973307
PUBMED
REFERENCE
2 (bases 1 to 2110)
Grauso, M. and Sattelle, D.B.
Direct Submission
Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional
Genetics Unit, University of Oxford, South Park Road, Oxford OX1
3QX, UK
FEATURES
source
1..2110
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="30D1"
/dev_stage="embryo"
1..2110
/gene="nAcRalpha-30D"
379..1950
/gene="nAcRalpha-30D"
/note="ion channel; neurotransmitter transmembrane
receptor; alternatively spliced; contains exons 3a, 8a and
8b; contains two repeats of the transmembrane TM2 region"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalpha6 subunit
variant type V"
/protein_id="AA013396.1"
/db_xref="GI:20152853"
/translation="MDSPLPASLSLFLVLLIPLAIKESCCQPHKRLNHLSTYNTL
BRPVANSEPELVKFLTQOI IDVDEKNOILTNAWLNLEWNDVNLWNTEYEGGVK
DLRTTPNKLKPDVLMYNSADGPDVTHTNVKNGSCLYVPPGIFKSTCKMDITW
FPDQHCMEKFGWYDCNQDLVNSDGDGLSDFTNGEWILLAMPKNTIVVA
CCPEYVDITTIQIRRTLYFFNLIVPCVLISSMALLGFTLPDPSGEKILGVTIL
LSLTVFLNVAESMPTSDAVPLIGTILLSLTVFLNVAETLPQVSDAIPULGYFN
CTMFVASSVVLTVVLMNHRTADIHMPPIKISVFLQWLPWILMRMGPGRKTRKT
ILLSNRMKELKERSKSLLANLVIDDDFRHTISGQTGSAIGSASFGRPTTVEHH
TAJGNHKDLHLILKELQFITARMKADDEALIGDWKFAAMVVDRECLIVPTLTII
ATVVLSSAPHILVQ"
846
/gene="nAcRalpha-30D"
/note="results in methionine to isoleucine substitution;
compared to the sequence deposited in GenBank Accession
Number AE003626"
/replace="a"
1401
/gene="nAcRalpha-30D"
/note="Compared to the sequence deposited in GenBank
Accession Number AE003626"
/replace="c"
1647
variation
variation
variation

TITLE Direct Submission
JOURNAL Submitted (19-APR-1999) ZF-BTB, Bayer AG, Bldg. Q 18,, Leverkusen
51368, Germany

FEATURES
source Location/Qualifiers
1..3629
/organism="Heliothis virescens"
/mol_type="mRNA"
/db_xref="taxon:7102"
335..1825
/note="hvnachra7-1"
/codon_start=1
/product="putative nicotinic acetylcholine receptor alpha
7-1 subunit"
/protein_id="AAD32697.1"
/db_xref="GI:4895005"
translation="MGGARRSHLAAPAGILLLLCLLWPRGARQGVYHEKRILLHLLDH
YVLERPVNESDPLQLSFLGLTLMQIIDVDEKQLLITNWLKLEWDMNLRWNTSDF
GGVKDLRVPRLHRLMNSADEGFDSTPTNVVVRNNGSLYVPPGIFKSTCKI
DIYFPDDQRCMEKFSWTYDGLDLODEGGDISFVINGEWELIIVPGRKNE
IYVNCPEYIDITFAVIRKTLTYFNLIVPCVLIASMALLGFTLPDPSGEKLSIG
VTLLSLTVELNVAETMPATSDAVPLGLTYFNCIMFWVASVSVSTLLILNYHHRHAD
THEMSDIWRCVELVWLPVLRMSRPSGATPPPARVPPDLRLRERSKSLIANVLD
IDDFRHPQAOQPCCRYYRGEBENGAGLAHSCFQVDYELSLILKEIRVITDQMRKD
DEDADISRWDKFAFNVDRLLIIFTLTIATLAVLLSAPHINVS"

ORIGIN
Query Match 38.3%; Score 575; DB 3; Length 3629;
Best Local Similarity 64.6%; Pred. No. 1.1e-141;
Matches 929; Conservative 0; Mismatches 460; Indels 48; Gaps 3;

Qy 64 CACGAGAGAGACTCTGAAACCGCTGTGTCGCGAAGTACAAACCCCTGGAGCGACCGGTG 123
Db 431 CACGAGAGCGCTACTGCACCACTATTGGACCACTACAAAGTACTGGAGAGCGCGTC 490

Qy 124 GCCAACAGAGCGAACCGCTAGAGTCAGGTTCCGCTTGACCTTGCAGCAATCATTTGAC 183
Db 491 GTCACAGAGCGACCGCTGCAGCTCTCTTCGCGCTCAGCTCATGTCAGATCATGAC 550

Qy 184 GTGACAGAGAGAACTCACTATTATACCAATATATGGCTGTGCTGGAGTGAATGAC 243
Db 551 GTGACAGAGAACCGAGCTTTTAAACAAACATCTGGCTTAAACTAGAGTGAATGAT 610

Qy 244 TAAACCTGAGTGGAGCGACGAGGATATGGCGGGGTCAAGGACCTCAGGATCAGCGCC 303
Db 611 ATGAACCTGAGTGGAGCACTTCAGATTTCCGCGGGGTCAAGATTTAAGAGTGCACCC 670

Qy 304 AACAGTGTGGAAGCGGAGCTCTTATGATATATAGTCTGACGAGGGTTTGACGG 363
Db 671 CACAGACTATGGAACACAGAGCTCTTATGATACACAGCGCGGAGAGGGTTTCGACAG 730

Qy 364 ACCTACAGAGCAACAGTGGTGTGAGAGCGCGGCGAGTTGCCCTGTACGTGCCACCTGGC 423
Db 731 ACGTATCCAGCAACGTGGTGTGCGGAACAAAGCTCGTCTGTACGTGCCCGCGGC 790

Qy 424 ATATTCAAGAGCAGATGCAAGATGGAACATCGCGTGGTTTCCCTTCGAGACCAACATGT 483
Db 791 ATCTTCAAGAGCAGCTGCAAGATGCAATCATCACTCTGGTTCCCTTCGAGCAACAGATG 850

Qy 484 GATATGAAGTTCGTAGCTGACATATGACGCGCAATCAGTTGGATCTGGTCTTAAAGAT 543
Db 851 GAGATGAAGTTGGAGCTGAGCTTATGATGTTTATCAGTTGGATCTTAACTACAGAT 910

Qy 544 GAGGAGCGCGGATCTATCGAGCTTATCAACAAATGGGAGTGGTATCTTAAATAGGAATG 603
Db 911 GAAGGGGCGGAGATATAAGCAGTTTGTACGAATGCGAATGGAGTTAATAGGATG 970

Qy 604 CCAGGCAAAAGAACACAAATACATACCGGTGTCGCCCGAGCCCTACGTGGAGTCAC 663
Db 971 CCCGGCAAGCGCAACAGATCTACTACAACTGTTGTCCGGAGCCATACATGACATCAG 1030

Qy 664 TTCACCATCATGATGAAGAGACGCTTGTACTACTTCTTCAACCTGATCGTCCCGTGC 723
Db 1031 TTTCCGGTGTGATCCGAGGAACGCTCTACTACTTCTTCAATCTGATCGTCCCGTGC 1090

Qy 724 GTGCTGATCTCATCGATGGCACTCCTCGGCTTTCACACTGCCACGAGACTCCGGAGAGAA 783
Db 1091 GTGCTCATCGCTCCATGGCTCTATTGGGGTTTACCTTGGCTCAGACTCCGGAGAAAG 1150

Qy 784 CTCACACTTGGAGTCACTATTCTTCTATCGCTGACGGTGTTCCTCAACCTGGTACCGAG 843
Db 1151 TTGCTTTAGTGTGACGATATTACTGCTGTGACGGTGTTCCTCAACATGGTGGCGAG 1210

Qy 844 ACCCTGCCACAGGTCTCCGACGCTATCCCTGTGTAGGGAGTACTTCAATTTGATCATG 903
Db 1211 ACGATGCCAGGACGTCGGACCGCTGCTCCCTTGTCTCGGCACTTCTCAACTGTCATG 1270

Qy 904 TTCATGTTAGTGTGCTGTGGTACTGACTGTGTGTGTACTCAATTAACCACTCAACA 963
Db 1271 TTCATGTTGGCTTCTCCGTCGCTCTCCACCATCTGATCTCTCAACTACCACCGGCAC 1330

Qy 964 GCTGATATCATGAATGCCACAGTGATGATAAATCAGTATTTCTTACAAATGGTTGCGATG 1023
Db 1331 GCAGACACTCAGCAATGAGTGTGATGTTGCGTGTTCCTTTATTTGGCTGCGGTG 1390

Qy 1024 ATACTGGAATGTCCAGGCGCAGGAGAAAGATCACACAGGAAGACTATAATGATGAACACG 1083
Db 1391 GTGCTGCGATGTACGCGCGGCTCGCGG---ACGACGCGCGCGCGCGGCTACCT 1447

Qy 1084 AGGATGAGGAGCTGGAACCTGAAGGAGAGGTCTGTCGAAGTCTTGTGCGGAAATGTTCTA 1143
Db 1448 CCGCGCGCGACCTGGAGCTGCGGAGCGCTCTCTCCAAGTCGCTCTTAGCGAACGTGCTC 1507

Qy 1144 GATATGATGATGATCTTACAGACGCGCTTCGCGCTCTTAACTGCTGCTGCTGCGCGG 1203
Db 1508 GACATGATGACGACTTCCGCCACCCGCAAGCGAGCAG-----CCGCAATGCTGCGG 1560

Qy 1204 AATTTGGGACCTGCGTCTCAATATCCGACGAGATTTCCGTCGCTGCTGCTGCTGCTCG 1263
Db 1561 ATACTACAGGGGGGTGAGGAGAAATGCGCGGGTTGCGCGCGCACAGTTGCTTCG---- 1616

Qy 1264 TCCACGATGAAGACGTTGGCGGGCTGGGTAGCCACCATCCGAGCTGCACTCATATA 1323
Db 1617 -----GTGTCGACTACGAGCTCTCCCTCAT 1642

Qy 1324 CTGAGAGCTGCGATTCATCAGCGCGAGATGAAGAGGCTGATGAGGAGCGGAGCTG 1383
Db 1643 CTGAAGAGATTAGAGTCATCAGATCAGATGCGCAAGGACGACGAGATGCGGACAT 1702

Qy 1384 ATCAGCACTGGAAGTTTGTGCGATGCTGTTGATAGTTTTCCTGTTGCTGTTTACA 1443
Db 1703 TCGCGCACTGGAAGTTGCGCGCATGCTGCTGAGACAGCTGTCCTTATTTTACC 1762

Qy 1444 CTTTTCACATCATCGGACAGTGTCTCTGTTATCGGACCGCATATCATCGT 1500
Db 1763 CTGTTCAATCATCGCCACGCTAGCGGTGCTGCTGTCGCGGCGCACACATCATGTTG 1819

RESULT 12
AX009612
LOCUS 3700 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 3 from Patent EP0962528.
ACCESSION AX009612
VERSION AX009612.1 GI:9996844
KEYWORDS
SOURCE Heliothis virescens (tobacco budworm)
ORGANISM Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE 1
AUTHORS Adamczewski M.D., Schulte, T.D. and Oellers, N.D.
TITLE Nucleic acids encoding acetylcholin-receptor subunits from insects
JOURNAL Patent: EP 0962528-A 3 08-DEC-1999;
BAYER AG (DE)
FEATURES
source Location/Qualifiers
1..3700

/organism="Heliothis virescens"	
/mol_type="unassigned DNA"	
/db_xref="taxon:7102"	
335..1825	
/note="unassigned protein product"	
/codon_start=1	
/protein_id="CAC07500.1"	
/db_xref="GI:9996845"	
/db_xref="REMRMBU: CAC07500"	
/translation="MGRRARRSHLAAPAGLALLLLCLLWPRGARGCYHEKRLHLHLLDH YVLERPVVNESDPLQSLFGLTLMQIDVDEKNQLLITNWLKLENDNMNLRWNTSDF GVKDLRPFDRORCEKMSWTYDGLDLOLQDEGGDISSFVINGEWELIIVPGKENE DITWPFDRORCEKMSWTYDGLDLOLQDEGGDISSFVINGEWELIIVPGKENE IYVNCPEPIDITFAVIRKLYFFNLIIVPCULIASMALLGFLPLPDSSEKLSLG VTLLSLTVLNMVATSDAVLELGTTFYENFMFVASSVSTILLINLHHRHAD THEMSDWIRCVLPWLMRSPGSAITPPRPVPPDLERLRSKSLANLVID IDDFHPQAPQPCRYRGGENGAGLAHSCFVGYELSLILKEIRVITDQMERK DEDADISRDKWFAAMVVDRLCLLIIFLFTIATLAVLLSAPHIMVS"	
ORIGIN	
Query Match 38.3%; Score 575; DB 6; Length 3700;	
Best Local Similarity 64.6%; Pred. No. 1.1e-141;	
Matches 929; Conservative 0; Mismatches 460; Indels 48; Gaps 3;	
QY 64 CACGAGAAGAGACTCTCGAACGGTGTGCTGGCGAACTACAAACCCCTGGAGCGACCGGTG 123	Db 1151 TTGCTTTTAGGTGTGACGATATTACTGTCTGACGGTGTCTCAACATGGTGGCGAG 1210
Db 431 CACGAGAAGCGGTACTGACCACTATTGGACCACTACAACTGCTGGAGAGCGCGTC 490	QY 844 ACCCTGCCACAGGTCTCCGACGCTATCCCCCTGTAGGACGCTACTTCAATTCATCATG 903
QY 124 GCCAAGCAGAGCGAAACCGCTAGAGGTGAGTTCGGCTTGACCTTGCAGCAAATCATTTGAC 183	Db 1211 ACGATGCCAGGACGTCGAGCGCGTCTGCTCGGACCTACTTCAATTCATCATG 1270
Db 491 GTCAACGAGAGCGACCGCTGCGAGCTCTCTTCCGCTCAGCTCATGACGATCATCGAC 550	QY 904 TTCATGGTAGCGTCTGTGTGTACTGACTGTGGTGTGTTACTCAATTTACCAACATCGACA 963
QY 184 GTGGAGAGAGAAATCACTATTAAACCAATATATGCTGTGCTGGTGGAGTGAATGAC 243	Db 1271 TTCATGGTGGTCTCTCCGCTGCTCCACCATATGATCTCAATTCATTCACCAACCGGAC 1330
Db 551 GTGGAGAGAGAAACCACTTTTAAACCAACATCTGGCTAAACCTAGAGTGAATGAT 610	QY 964 GCTGATATATCATGAATGCAAGTGCAGAGTGAATAAATCAGTATTCCTCAATGTTGGCATGG 1023
QY 244 TACAACCTGAGTGGAGACGACGAGTATGGCGGCTCAAGACCTCAGGATCAGCCC 303	Db 1331 GAGACACATCAGAAATGATGATGATTCGTTGGTGTCTTATTTGGCTGGCGTGG 1390
Db 611 ATGAACTTGGAGTGGAAACCTTCAAGATTTGGCGGGGTCAAGATTTAAGAGTGCACCC 670	QY 1024 ATACTCGAATGTCGAGGCGAGGAAAGATCACAGGAAGACTATAATGATGAACAG 1083
QY 304 AACAAAGTTGGAAAGCGGACGCTTATATATAATAGTCTGACGAGGGTTTTGACGGG 363	Db 1391 GTGCTCGCATGTACAGCGCGGCTCGCG--ACGACGCGCGCGCGCGCTACCT 1447
Db 671 CACAGACTATGAAACACAGAGCTCTTATGACAGCGCGGACGAGGTTGACAGC 730	QY 1084 AGGATGAGGAGCTGGAATGGAAGGAGGTGCTGCAAGTCTTCTGCTGGCAATGTTCTA 1143
QY 364 ACCTACAGACCAACGTGGTGTGACAGCGCGGCGAGTTCCTGTACGTGCCACTGGC 423	Db 1448 CGCCGCCCGGACCTGGAGCTGGCGAGCGCTCTCAAGTCTCTTAGCGAAACGTGCTC 1507
Db 731 ACGTATCCACGAACGTGGTGTGCGGAAACAACGGCTCGTGTCTGACGTGCGCGCGCGC 790	QY 1144 GATATTGATGATGACTTTCAGACACGCGCTCTCCGCTCTTAACAGTACTGCTCGACGGG 1203
QY 424 ATATTCAAGACACATGCAAGATGGACATCGCTGCTGCTTCCCTTCGACGACCAACTGT 483	Db 1508 GACATCGATGACGACTTCCGCCACCGCAAGCGCAGCAG-----CCGCAATGCTGCCG 1560
Db 791 ATCTTCAAGAGCACTGCAAGATCGACATCACTGTTCCCTTCCGACGACCAAGATGC 850	QY 1204 AATTTGGACCTGGGTGCTCAATATTCGACAGGATTTCCGTGGTGTGCTGGTTCGTCGG 1263
QY 484 GATATGAAGTTGCGTGTGACATATGACGGAATCAGTTGGATCTGTGCTGCTAAAGAT 543	Db 1561 ATACTACAGGGGGGTGAGAGAAATGGCGGGGTTGGCGGCGCAGATGCTTTCG--- 1616
Db 851 GAGATGAAGTTTGGCAGCTGCACTTATGATGTTATCAGTTGGATCTPACAACTACAGAT 910	QY 1264 TCCACATGGAAGACGTGGGCGGCTGGGTAGCCACATCGGAGCTGCACCTCAT 1323
QY 544 GAGCAGCGCGGATCTATCGGACTTCAACAAATGGGAGTGTGTTCTAATAGGATG 603	Db 1617 -----GTGCGACTACGAGCTTCCCTCAIT 1642
Db 911 GAAGGGGGCGAGATATAAGCAGTTTGTACCAATGGCGAATGGGAGTTAATAGGAGTC 970	QY 1324 CTGAGAGAGCTGCAGTTTCATCAGGCGAGGATGAAGAAGCTGATCAGGAAGCGAGCTG 1383
QY 604 CCAGGAAAAAGAACACAATACATACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663	Db 1643 CTGAGAGATTAGATGATCATCAGATCAGATGCGCAAGGACGACGAGATGCGGACAT 1702
Db 971 CCAGGCAAGCAAGAGATCTACTCACTGTTGCTGCGAGGCAATCATCATGATCAGC 1030	QY 1384 ATCAGCAGCTGGAAGTTTCTGCGATGTTGTTGATAGTGTTCCTGTTGCTGTTGTTGTTGTT 1443
QY 664 TTCAACATCATATAAGAACGAACTTGTACTTCTTCAACCTGATCGTCCCGTGC 723	Db 1703 TCGCGGACTGGAAGTTGCGCGCATGTTGCTGACAGACTGTGCTTATTTATCTTTACC 1762
Db 1031 TTTGGGTGTGTATCCGAGGAAACGCTCTACTTCTTCTCAATCTGATCGTGCCTGC 1090	QY 1444 CTTTTCACATCATCGGACAGTAGCTGTCTGTTATCGGACCGGATATCATCTG 1500
QY 724 GTGCTGATCTATGATGGCACTCTCGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 783	Db 1763 CTGTTCAATCATATCGGCGAGTGGCTGTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1819
Db 1091 GTGCTCATCGCTCCATGGCTTATTTGGGTTTCACTTGGCTTCCAGACTCCGAGAGAAAG 1150	RESULT 13
QY 784 CTCACACTTGGAGTCACTATTTCTTCTATCGTGTGAGGCTGTTCTTCACTGAGCGAG 843	E58347
	LOCUS
	DEFINITION
	Nucleic acid encoding insect acetyl choline receptor subunit.
	ACCESSION
	E58347.1
	VERSION
	GI:13019346
	KEYWORDS
	JP 2000023680-A/2.
	SOURCE
	Heliothis virescens (tobacco budworm)
	ORGANISM
	Heliothis virescens
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuidae; Noctuidae; Heliothinae; Heliothis.
	REFERENCE
	1 (bases 1 to 3701)
	Martin, A., Nadja, E. and Thomas, S.
	Authors
	Nucleic acid encoding insect acetyl choline receptor subunit
	Title
	Patent: JP 2000023680-A 2 25-JAN-2000;
	Journal
	BAYER AG
	COMMENT
	OS Heliothis virescens
	PN JP 2000023680-A/2
	PD 25-JAN-2000
	PF 26-APR-1999 JP 1999118159
	PR 04-MAY-1998 DE 19819829.9
	PI MARTIN ADAMUTSUSUKI, NADJA ERASU, THOMAS SCHULTE PC
	C12N15/09, A01K67/03, C07K14/705, C07K16/28, C12N1/21, C12N5/10, PC
	C12Q1/68,

JOURNAL	Genetics 160 (4), 1519-1533 (2002)	
MEDLINE	21969411	
PUBMED	11973307	
REFERENCE	2 (bases 1 to 2907)	
AUTHORS	Grauso, M. and Sattelle, D. B.	
TITLE	Direct Submission	
JOURNAL	Submitted (27-MAY-2000) Human Anatomy and Genetics, MRC-FGU,	
FEATURES	University of Oxford, South Park Road, Oxford OX1 3QX, UK	
source	Location/Qualifiers	
	1..2907	
	/organism="Drosophila melanogaster"	
	/mol_type="mRNA"	
	/db_xref="taxon:7227"	
	/chromosome="2"	
	/map="2L; 34E4-5"	
	/dev_stage="embryo"	
gene	1..2907	
	/gene="nAcRalpha-34E"	
	/allele="A"	
	289..2712	
CDS	/gene="nAcRalpha-34E"	
	/note="ion channel; neurotransmitter transmembrane receptor"	
	/codon_start=1	
	/product="nicotinic acetylcholine receptor Dalphas subunit"	
	/protein_id="AAAI3390.1"	
	/db_xref="GI:20152840"	
	/translation="MKNAOLKLTVEDDDELWLAHLAHCNSFSSSSSTRITSSNQRH NQULTLQPSLSLTKHNSIAEQHNSQQQEPASKEDVANHGKSNQQTHLQQLDSS NMUSPKTAAATAGDEATTQOPTNIRLCARQRRLRRKRKPATNETD1KKQOOL SMPFKRTADTYSPTAATSCPTATMOCRASDNFESIPISRHRVSTATFAWVLH VLVQLVSLQWOLHWQVRSVLLFRIAASTAFISYLSGFAOLKNSSSSSSSSNS NSSTQLNGKHSWIFLLIYLNLSAKVCLAGVHEKRLHLDDLDPYNTLERPVINES DPLQSLGLTLMOLIDVDEKNQLLVTVNWLKLEWDMNLEWNSDYGVXDLRIPHR IWRPDLMYNSAGEFGDTQTNVVRNNGSCLYPPGIFKSTCKIDITWFFPDDQRC EMPKGSWTFQDQLQDETGDIDSYVLNGEWELLGVPGRNEIYVNCPEPYD ITFAIIRRTLYFFNLIIPCVLIASMALLGFTLPDPSGEKLSGLVTLILSLTVFLN MVAETMPATDAPLVLCIFNCIMFVASSVSVTILINVHHRNADTHEMSEWIRVF LCLWPLTMRSPGRPLILEPPTPCSDTSERKHOILSDVELKERSKSLIANVLDI DDPFRNCRPMTGGTLPHNPAFYRTYGGDDGSGIGPISGTMPDVAVTHTCIKSST EYELGLIKELIRFITDQKRKDECDNIANDKFAAMVVDRLCILIIFTMFALLTAVL LSAPHIIVS"	
variation	373	/gene="nAcRalpha-34E"
		/note="compared to B allele"
variation	375	/replace="a"
		/gene="nAcRalpha-34E"
		/note="results in phenylalanine to isoleucine substitution; compared to B allele"
variation	827	/replace="c"
		/gene="nAcRalpha-34E"
		/note="results in threonine to isoleucine substitution; compared to B allele"
variation	1079..1080	/replace="t"
		/gene="nAcRalpha-34E"
		/note="results in lysine to arginine substitution; compared to B allele"
variation	1081..1086	/replace="gg"
		/gene="nAcRalpha-34E"
		/note="results in deletion of asparagine and serine; compared to B allele"
variation	1119^1120	/replace="n"
		/gene="nAcRalpha-34E"
		/note="results in serine, glycine, and asparagine insertion; compared to B allele"
variation	1121	/replace="agcggcaac"

/gene="nAcRalpha-34E"

/note="results in asparagine to serine substitution; compared to B allele"

/replace="g"

ORIGIN	Query Match	34.7%;	Score 521.6;	DB 3;	Length 2907;
	Best Local Similarity	61.8%;	Pred. No. 1.9e-127;		
	Matches 924;	Conservative 0;	Mismatches 519;	Indels 51;	Gaps 4;
Qy	57	AGGTCCTCAGAGAAGAGACTCTGAAACGGGTTCCTGCGGAACATACAAACCCCTGGAGCG	116		
Db	1212	AGGATATCATGAAAAGAGACTGTTACACGATCTTTTGGATCCTTATAATACACTAGAACG	1271		
Qy	117	ACCGTGGCCAAACGAGAGCGAACCGCTAGAGTCAGGTCGGCTTGACCTTGCAGCAAT	176		
Db	1272	TCCGGTTCTCAATGAATCGGACCCGTTACAATTAAGCTTTGGTTAACTTTAATGCAAT	1331		
Qy	177	CATTGACGTGGACGAGAGAAGAACTCAACTATTATAACCAATATATGGCTGTCTTGGAGTG	236		
Db	1332	TATCGATGTGGACGAGAAAATCAATTGCTAGTCACTAAATGTGTGGTTAAACTGGAGTG	1391		
Qy	237	GAATGACTACAACTGAGGTGGAAACGACGAGAGTATGCGGGGTCAAGGACCTCAGAT	296		
Db	1392	GAACGACATGAATCTCCGCTGGAAACCTCCGACTATGGCGAGTTAAGGATCTCGCAAT	1451		
Qy	297	CACGCCCAACAAGTTGTGGAAGCGGACGCTCTTATGTATAATAGTGTGACGAGGGTTT	356		
Db	1452	ACCGCCCATCGCATCTGGAAAGCCGACGCTGCTGATGACAACTGCGATGAGGGATT	1511		
Qy	357	TGACGGACCTTACCAGACCAACGCTGTGTGTCAGAAAGCGGCGAGTTGCCTGTACGTGCC	416		
Db	1512	TGACGGCACCTACCAGACGAACGCTGTGTGTCGGAACAAACGGCTGTCTATACGTTCC	1571		
Qy	417	ACCTGGCATATTCAAGAGCAGATCGAAGATGGACATCGGTGTTTCCCTTCGACGACCA	476		
Db	1572	GCGGGGATCTTCAAGTCGACGCTGCAAGATCGACATCACTGCTTCCCTTCGATGACCA	1631		
Qy	477	ACACTGTGATGAAGTTCCGTAGCTGGACATATAGCGCAATCAGTTGGATCTGGTGCT	536		
Db	1632	GCGGTGCGAGATGAAGTTCCGCGAGTTGGACCTACGACGATTCAGCTGGATTTACAATT	1691		
Qy	537	AAAAGATGAGCGAGCGCGGATCTATCGGACTTCTATCAACAAATGGGAGTGGTACTAAT	596		
Db	1692	ACAAATGAAACTGGCGGTGATATCAGCAGTACGTCTCAACGCGAGTGGGAATCT	1751		
Qy	597	AGGAATGCCAGGCAAAAAGAACACATAACATACGCGTGTGCCCCGAGCCCTACGTGGA	656		
Db	1752	GGGTGTGCCCGGCAACGTAACGAGATCTATTACAACTGCTGCCCGGAACCTATATAGA	1811		
Qy	657	CGTCACCTTCACCATCATGATAAGAACGAAACCTTGTACTACTTCTTCAACCTGATCGT	716		
Db	1812	CATCACCTTCGCCATCATTCGCCCGAACACACTGTACTATTCTTCAACCTGATCAT	1871		
Qy	717	CCGTCGCTGCTCATCTCATCGATGGCACTCTCGGCTTCACACTGCCACAGACTCCGG	776		
Db	1872	ACCTTGTGTACTGATTCGCTCCATGGCCTTGTCTGGATTCACTCCCGCAGATTCGG	1931		
Qy	777	AGAAAACCTCACACTTGGAGTCACTATTCTTATCGCTGACGGTGTTCCTCAACCTGGT	836		
Db	1932	TGAAAAATTATCGCTGGGTGTTACCATCTTGTCTCTCGCTGACCGGTGTTCTGAATATGGT	1991		
Qy	837	AGCGGAGACCTTGCACAGGCTCTCCGACGCTATCCCCCTGTAGGGACGACTTCAATG	896		
Db	1992	TGCCGAGACAAATGCCGCTACTTCCGATGGGTGCCATCTGCTGGGTACATATTCAATTG	2051		
Qy	897	CATCATGTTTCATGGTAGCGTCTGCTGTGTGACTGTGACTGTGGTGTACTCAATTACACCA	956		
Db	2052	CATAAATGTTTATGGTAGCTTTCATCCGTTGTGTCAACGATTTTAATAATAAATATATCA	2111		
Qy	957	TCCGAACGCTGATATACATGAATGCCACAGTGGATAAATCAGTATTCCTACATCGTT	1016		
Db	2112	TGAAATGCTGATACGACGCAAAATGTCGGAATGATACGCAATCGTGTGCTGCTGCTGCT	2171		

QY	1017	GCATCGTACTCGGAATGTCGAGGCCAGGAAGAGATCACCAGGAG-----	1065
Db	2172	GCATGGATATTGGAATGAGTCGCCGAGGACGCGCTGATCCTAGAGTTCCCGACCA	2231
QY	1066	-----ACTATAATGATGAACACGAGGATGAGGAGCTGGAACGTAA	1106
Db	2232	GCCTGTTCCGACACATCCTCCGAGCGGAGCACCAGATCTCTCCGACGTTGAGCTGAA	2291
QY	1107	GGAGAGTCTGTCGAAGTCTTGTGCGGAATGTTCTAGATATTGATGATGACTTCAGACA	1166
Db	2292	AGAGCGCTCTCGAAATCGCTGTGCGCAACGTACTAGACATCGATGACTTCCGGCA	2351
QY	1167	CGGCCCTC-----CGCCTCCTAACAGTACTGCCTCGACCCGGGAATTTGGGACCT	1215
Db	2352	CAATGTGCGCCCATGACGCCCGCGGAACACTGCCACAACCCCGCTTTCTATCGCAC	2411
QY	1216	GGGTGCTCAATATTCCG-CACGGATTTCGGTCGGTCGTCGGTCCGTCACGATGGA	1274
Db	2412	GGTTATGGACAAGGCGACGATGGCAGCATTTGGCCCAATTGGCAGCACCCGGAATGCCGA	2471
QY	1275	AGACGTGGCGGGCGGTGGTAGCCA-----CCATCGCGAGCTGCACCTCATCT	1325
Db	2472	TGCGGTACCCCATCATACGTGTCATCAAAATCATCAACTGAATATAGTTTAACTTT	2531
QY	1326	GACAGAGCTGCAGTTTCATCAGCGCCAGGATGAAGAGGCTGATGAGGAAGCCGAGCTGAT	1385
Db	2532	AAAGGAAATTCGCTTTATAAATGATCAGCTACGTAAGATGACGAGTGCAATGACATTCG	2591
QY	1386	CAGCGACTGGAAGTTTGCTCGGATGGTTGTTGATAGGTTTTCGCTGTTTCGTTCACACT	1445
Db	2592	CAATGATTGGAATTTGCACCTATGGTCGTTGACAGACTGTGCTTATCATATTCACAAT	2651
QY	1446	TTTCACAAATCATCGGACAGTAGCTTCCTGTTATCGGCACCGCATATCATCGT	1499
Db	2652	GTTGCAATATTAGCCACAATAGCTGTACTACTATCATCAGCACCAATATTATTGT	2705

Search completed: May 8, 2004, 02:41:32
 Job time : 5963.59 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 10:08:00 ; Search time 604.489 Seconds

(without alignments)
10562.710 Million cell updates/sec

Title: US-09-303-232-5_COPY_95_1597

Perfect score: 1503

Sequence: 1 atggccctcatgtggcgc.....caccgcatacatcgtgcaa 1503

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1503	100.0	3109	3	Aaz24477 H. viresc
2	575.6	38.3	1540	4	AbL07231 Drosophil
3	575	38.3	3700	3	Aaz24476 H. viresc
4	430.6	28.6	936	4	AbL13733 Drosophil
5	411.4	27.4	2886	3	Aaz24475 D. melano
6	377.2	25.1	1509	3	Aac58395 Human PRO
7	374	24.9	1509	4	Aac90380 Wild-type
8	374	24.9	1876	2	Aat48239 Neuronal
9	374	24.9	1876	8	ADA10864
10	373.6	24.9	1530	2	AAV44687 V274T var
11	373.6	24.9	1876	2	AAV12197
12	373.6	24.9	1876	6	ABSS4875
13	373.6	24.9	1876	6	ABV73248
14	372.4	24.8	1509	4	AAC90385
15	372.4	24.8	1509	4	AAC90386
16	370.8	24.7	1509	4	AAC90387
17	363.2	24.2	1964	6	ABZ11298
18	342.4	22.8	2769	2	AAT59196
19	326.8	21.7	2101	2	AAT59197
20	322.8	21.5	1869	6	AAAL45873
21	322.8	21.5	1869	6	ABL54794 Modified
22	304.4	20.3	2082	2	AAT59527
23	304.4	20.3	2752	9	ADB78668 Human nic

24	304.4	20.3	2752	9	ADB78671	Human nic
25	304.4	20.3	2752	9	ADB78670	Human nic
26	304.4	20.3	2752	9	ADB78661	Human nic
27	304.4	20.3	2752	9	ADB78672	Human nic
28	304.4	20.3	2752	9	ADB78663	Human nic
29	304.4	20.3	2752	9	ADB78669	Human nic
30	304.4	20.3	2752	9	ADB78662	Human nic
31	304.4	20.3	2752	9	ADB78673	Human nic
32	304.4	20.3	3343	9	ADB85024	Farnesyl
33	304.4	20.3	3496	2	AAT48237	Neuronal
34	304.4	20.3	3496	8	ADA10858	Human neu
35	303.8	20.2	2666	9	ADC71168	Human neu
36	302.8	20.1	1809	5	AAS91552	DNA encod
37	302.8	20.1	2082	2	AAT59528	Alpha4 su
38	302.8	20.1	2752	9	ADB78660	Human nic
39	302.2	20.1	1908	2	AAT48236	Neuronal
40	302.2	20.1	1908	8	ADA10856	Human neu
41	302.2	20.1	2277	2	AAV12199	Human neu
42	302.2	20.1	2277	2	AAQ90387	Alpha 2 s
43	302.2	20.1	2277	6	ABS54870	Human neu
44	302.2	20.1	2277	6	ABV73243	Human neu
45	302.2	20.1	2664	2	AAT48235	Neuronal

ALIGNMENTS

RESULT 1

AAZ24477

ID AAZ24477 standard; cdna to mRNA; 3109 BP.

XX

AC AAZ24477;

XX

DT 17-FEB-2000 (first entry)

XX

H. virescens acetyl-choline receptor DNA from clone Hva7-2.

XX

Acetyl-choline receptor; nicotinic; insect; insecticide; screening; neurotransmission; plant protection agent; conductance; AChR; ds.

XX

OS Heliothis virescens.

XX

PN DB19819829-Al.

XX

PD 11-NOV-1999.

XX

PF 04-MAY-1998; 98DE-01019829.

XX

PR 04-MAY-1998; 98DE-01019829.

XX

PA (FARB) BAYER AG.

XX

PI Adamczewski M, Oellers N, Schulte T;

XX

WPI: 2000-014207/02.

DR P-PSDB; AAY50816.

XX

New nucleic acid encoding a nicotinic acetylcholine receptor from insects, used to identify potential insecticides.

PT

Claim 1a; Page 19-22; 26pp; German.

XX

This invention describes a novel nucleic acid (NA) encoding a nicotinic acetyl-choline receptor (I) from insects which can be used as an insecticide. Inhibitors of (I) interfere with neurotransmission. (I) (also vectors containing it, its regulatory regions, and antibodies directed against (I)-encoded proteins) are used to screen for: (a) plant protection agents that alter conductance of AChR, potentially useful as insecticides, or (b) genes which encode polypeptides that are involved in formation of functionally related AChR in insects. (I) are also used to isolate and characterize the specified regulatory regions and for recombinant production of (II). This sequence encodes an acetyl-choline receptor isolated from *Heliothis virescens*

XX Drosophila melanogaster.
 OS WO200171042-A2.
 PN XX
 XX 27-SEP-2001.
 PD XX
 PF XX
 PE XX
 XX 23-MAR-2001; 2001WO-US009231.
 PR XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX PA
 PA (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 XX DR P-PSDB; ABB69630.
 DR XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX Claim 1; SEQ ID NO 35681; 2ipp + Sequence Listing; English.
 PS The invention relates to an isolated nucleic acid detection reagent
 XX capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from wifo at ftp.wifo.int/pub/published_pct_sequences
 XX SQ
 XX Sequence 936 BP; 238 A; 213 C; 239 G; 246 T; 0 U; 0 Other;
 Query Match 28.6%; Score 430.6; DB 4; Length 936;
 Best Local Similarity 68.5%; Pred. No. 2.4e-109;
 Matches 595; Conservative 0; Mismatches 274; Indels 0; Gaps 0;
 QY 19 GCCTTGGCGCTGTGGCTTGTGTCGCGGTATCGAGCAAGTCTCCACGAGAGAGACTC 78
 Db 67 GGCTGGGACTTTTAATTATGATACCGGCTTGTGCGGTGGACCCCATGAGAACGGCTA 126
 QY 79 CTGAACGGCTTGTGCGGAACCTACACACCTCGAGCGACCGGTGGCCCAACGAGCGAA 138
 Db 127 CTCACGCGCTTCTGGACAACTACAACAGCCTGGAGCGTCCGGTGCATGAATCCGAT 186
 QY 139 CGCTAGAGGTGAGTTCGGCTTGACCTTCAGCAAAATATTACCTGGACGAGAGAAAT 198
 Db 187 CCATTGCACTGAGCTTCGAGCTTAACACTCATGCAATATCATGATGTGACGAAAGAAAT 246
 QY 199 CAACTACTTATAACCAATATATGGCTGTGCTTGGAGTGAATGACTACAACTGAGGTGG 258
 Db 247 CAACTGCTTATACGATATTTGGCTCAATTGAATGGAACGATGATGATCTTCGATGG 306
 QY 259 AACGACGAGGATATGGCGGGTCAAGGACCTCAGGATCAGCCCAACAAAGTTGTGGAAG 318
 Db 307 AATTTCAGTGAAGTTCGGTGTGCGGGATCTCGGAATTCGCGCATCTCGCCTATGAAA 366
 QY 319 CCGGAGCTCTTATGATATGATGCTGAGGAGGTTTTCGCGGACCTTACGAGCAAC 378
 Db 367 CCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 426
 QY 379 GTGGTGGTCAGAGCGCGGCGAGTTCGCTGTAGTCCCACTCGCATATTTCAAGACCA 438
 Db 427 GTGGTGGTTCGCAATTAATGGAGCTGTCTGTACCTACCGCGCATATTTTATGTCACG 486
 QY 439 TGAAGATGAGATCCCGTGGTTCCTTCGACGACCAACACTGTGATATGAATTCGGT 498
 Db 487 TGAAGATGAGATTCAGTGGTTCCTTCGACGATGATGATGATGATGATGATGATGATGAT 546

QY 499 AGCTGACATATGACGGCAATCAGTTGGATCTGGTCTAAAAGATGAGGAGCGCGCAT 558
 Db 547 TCGTGGACCTTACGATGGGTTTCAGTTGGACCTTCAGTTGGACGAAAGCTGGTGGCGAC 606
 QY 559 CTATCGGACTTCATAACAATGGGAGTGGTATCTAATAGGAATGCCAGCAAAAGAAC 618
 Db 607 ATTCTAGCTTTATAACCAATGGGAATGGACTTGTAGGTGTGCCCGGTAAACGAAAT 666
 QY 619 ACAATAACATACGCTGCTGCCCGGAGCCCTACGTGGACGTCACCTTCCACCATCATGATA 678
 Db 667 GAAATCTACTATAATTTGCTGCCAGAACCTTATTGACATACATTCGCCATTTTGATA 726
 QY 679 AGAAGACGAACCTTGTAATCTTCTTCAACCTGATCGTCCCGTGGTGTGATCTCATCG 738
 Db 727 AGCGCAAAACGTTGTACTATTTTCAATCTGATTTGCGGTGCGTACTGATCGCGCTCC 786
 QY 739 ATGGCACTCTCGGCTTCACTGCGCCACGAGCTCCGAGAGAACTCACACTTGGATC 798
 Db 787 ATGGCACTCTGAGGTTTACCTGCCACGAGATCTGGTGAAAGCTTTTCGCTTGGAGTT 846
 QY 799 ACTATTCTTCTATCGCTGACGGTGTTCCTCAACCTGCTAGCCGAGACCTTGCACACAGTC 858
 Db 847 ACAATTCTATTATCGCTTACAGTCTTCTCAACATGCTGCGCGAACAATGCCGCGGACC 906
 QY 859 TCGACGCTATCCCTCTGTTAGGAGCTA 887
 Db 907 TCGATGCGGTACCGCTGCTCGGTAAGTA 935

RESULT 5
 AA224475
 ID AA224475 standard; cDNA to mRNA; 2886 BP.
 XX AC
 XX AA224475;
 XX DT 17-FEB-2000 (first entry)
 XX DE
 XX D. melanogaster acetyl-choline receptor DNA from clone Da7.
 XX KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
 XX KM neurotransmission; plant protection agent; conductance; AChR; ds.
 XX OS Drosophila melanogaster.
 XX FH Key Location/Qualifiers
 FT CDS 372..2684
 FT /*tag= a
 FT /*product= "acetyl choline receptor."
 XX DE19819829-A1.
 XX PD 11-NOV-1999.
 XX PF 04-MAY-1998; 98DE-01019829.
 XX PR 04-MAY-1998; 98DE-01019829.
 XX PA (FARB) BAYER AG.
 XX PI Adamczewski M, Oellers N, Schulte T;
 XX WPI; 2000-014207/02.
 DR P-PSDB; AAY50814.
 XX New nucleic acid encoding a nicotinic acetylcholine receptor from
 PT insects, used to identify potential insecticides.
 XX Claim 1a; Page 8-12; 26pp; German.
 CC This invention describes a novel nucleic acid (NA) encoding a nicotinic
 CC acetyl-choline receptor (I) from insects which can be used as an
 CC insecticide. Inhibitors of (I) interfere with neurotransmission. (I)

CC	(also vectors containing it, its regulatory regions, and antibodies
CC	directed against (I)-encoded proteins) are used to screen for: (a) plant
CC	protection agents that alter conductance of AChR, potentially useful as
CC	insecticides, or (b) genes which encode polypeptides that are involved in
CC	formation of functionally related AChR in insects. (I) are also used to
CC	isolate and characterize the specified regulatory regions and for
CC	recombinant production of (II). This sequence encodes an acetyl-choline
CC	receptor isolated from <i>Drosophila melanogaster</i>
XX	
SQ	Sequence 2886 BP; 942 A; 687 C; 640 G; 617 T; 0 U; 0 Other;
	Query Match 27.4%; Score 411.4; DB 3; Length 2886;
	Best Local Similarity 68.8; Pred. No. 8.e-104;
	Matches 565; Conservative 0; Mismatches 256; Indels 0; Gaps 0;
Qy	57 AGGTCCTCAGAGAAGACATCTCGAACGGGTTCCTGGCGAATCTACAACACCCTTGAGCG 116
Db	1295 AGGATATCATGAAGAAGACTGTTTACACGATCTTTTGGATCCTTATAATACACTAGAACG 1354
Qy	117 ACCGGTGGCCAAACGAGAGCGAACCGCTAGAGGTCAAGTTCGGCTTGCCTTGACGCAAAAT 176
Db	1355 TCCGGTTCTCAATGAATCGAGCCCGTTACAATTAAAGCTTTGGTTTAATTAATGCAAAAT 1414
Qy	177 CATTCAGTGGACGAGAAGAATCAACTACTTATAACCAATATATGGCTGTCTTGGAGTG 236
Db	1415 TATCGATGTGGACGAGAAAAATCAATTTGCTAGTCAC'TAA'TGTGTGTTAAAAC'TGGAGTG 1474
Qy	237 GAATGACTACAACTTAGGTTGGACGACAGCGAGTATGCCGGGTCAAGNACCTCAGGAT 296
Db	1475 GAACGACATGAATCTCCGCTGGAAACCTTCGCACTATGGCGGAGTTAAGGATCTGCCAAAT 1534
Qy	297 CAGCCCCAAACAAGTTGTGGAAAGCGGACGTCCTTTATGTATAATAGTGTGCACGAGGGTTT 356
Db	1535 ACCGCCGATCGCATCTGSAAGCGGACGTGCTGATGACAAAGTGGGATGAGGGATT 1594
Qy	357 TGACGGGAACCTACAGACCAACGTCGTGGTTCAGAAAGCGGCGGACAGTTCCTGTACTGTC 416
Db	1595 TGACGGCACCTTACAGACGAACGTCGTGGTGGGAAACAAGCGCTGTCTATACGTTC 1654
Qy	417 ACCTGGCATATTTCAAGAGCACATGCAAGATGGACATCGCGTGGTTTTCCCTTCGACGACCA 476
Db	1655 GC CGGGGATCTTCAAGTCGAGCTGCAAGATCGACATCACGCTGGTTCCCTTCGATGACCA 1714
Qy	477 ACATCTGTATATGAAGTTTCGTTAGCTGGACATATGACGGCAATCAGTTTGNATCTGGTCT 536
Db	1715 GCGGTGGAGATGAAGTTTCGACAGTTGGACCTACGACGGATTCACAGCTGGATTTACAATT 1774
Qy	537 AAAAGATGAGCGACGGCGCGATCTATCGGACTTCTATAACAATGGGAGTGGTATCTAAT 596
Db	1775 ACAAGATGAAC'TCGCGGTGATATCAGCAGTTACGTGCTCAACGGCGAGTGGGA'CTACT 1834
Qy	597 AGGAATGCGAGGCAAAAAGAACACAATAACAATACGCGTGTGCCCCGAGGCCCTACGTGGA 656
Db	1835 GGGTGTGCCCGGCAAAAGTAAACGAGATCTATTTAACTGCTGCCCGGAAACCCATATAGA 1894
Qy	657 CGTCACCTTCACCATCATGATTAAGAAGACGAACCTGTGTACTCTTCTCAACCTGATCGT 716
Db	1895 CATCACCTTCGCCATATCATCCGCCAGAGAACACTGTACTATTTCTTCAACCTGATCAT 1954
Qy	717 CCCGTGCGTGTGATCTCATCGATGGCACTCCTCGGCTTCACCTGCCACAGACTCCGG 776
Db	1955 ACCTTGTGTACTGATTCCTCCATGGCCTTGTCTCGGATTCACCTGCGCGAGATTCGGG 2014
Qy	777 AGAAGAACTCACATTCGAGTCACTATTCTTCTATCGTCAAGGTGTCCTCAACCTGGT 836
Db	2015 TGAANAATTTATCGCTGGGTGTACATCTTGTCTCTCGCTACCGGTGTTTCTGAATATGGT 2074
Qy	837 AGCGAGACCTTGCACAGGTCCTCCGAGCTATCCCCTGT 877
Db	2075 TGCCGAGACAAATGCCGGTCACTTCCGATGCGGTGCCATTGT 2115

RESULT 6

AAC58395
ID AAC58395 standard; cDNA; 1509 BP.
XX
AC AAC58395;
XX
DT 29-JAN-2001 (first entry)
XX
DE Human PRO2145 nucleotide sequence SEQ ID NO:76.
XX
KW Human; tumour; diagnosis; neoplastic disease; neoplastic cell growth;
KW proliferation; tumorigenesis; identification; cancer; cytostatic;
KW nontropic; neuroprotective; antiinflammatory; immunosuppressive;
KW immunostimulant; antiangiogenic; leukaemia; lymphoid malignancy;
KW neuronal disorder; glial disorder; astrocytal disorder; angiogenic;
KW hypothalamic disorder; glandular disorder; macrophagal disorder;
KW epithelial disorder; stromal disorder; blastocoelec disorder;
KW inflammatory disorder; immunologic disorder; ss.
XX
OS Homo sapiens.
XX
XX
XX WO2000053755-A2.
XX
XX 14-SEP-2000.
XX
XX 06-JAN-2000; 2000WO-US000376.
XX
XX 08-MAR-1999; 99WO-US005028.
XX 02-JUN-1999; 99WO-US012352.
XX 23-JUN-1999; 99US-0141037P.
XX 07-JUL-1999; 99US-0143048P.
XX 26-JUL-1999; 99US-0145698P.
XX 30-NOV-1999; 99WO-US028313.
XX 20-DEC-1999; 99WO-US030911.
XX 05-JAN-2000; 2000WO-US000219.
XX
XX (GETH) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hillan KJ, Roy MA;
XX Watanabe CK, Wood WI;
XX
XX WPI; 2000-572270/53.
XX
XX P-FSDE; AAB24086.
XX
XX Thirty PRO polynucleotides encoding PRO polypeptides, useful in the
XX treatment, diagnosis and prevention of cancer.
XX
XX Claim 50; Fig 57; 286pp; English.
XX
XX The present invention describes an isolated antibody that binds to one of
XX the human PRO proteins designated PRO212, PRO290, PRO341, PRO535, PRO619,
XX PRO717, PRO809, PRO830, PRO848, PRO943, PRO1005, PRO1009, PRO1025,
XX PRO1030, PRO1097, PRO1107, PRO1111, PRO1153, PRO1182, PRO1184, PRO1187,
XX PRO1281, PRO23, PRO39, PRO834, PRO1317, PRO1710, PRO2094, PRO2145 OR
XX PRO2198. PRO antagonists can be used to inhibit tumour cell growth. The
XX PRO polypeptides and nucleotides are useful in the treatment, diagnosis
XX and prevention of cancer. The antibodies and other anti-tumour compounds
XX maybe used to treat various conditions, including those characterised by
XX overexpression and/or activation of the amplified PRO genes. Exemplary
XX conditions or disorders to be treated with such antibodies and other
XX compounds include benign or malignant tumours (e.g., renal, liver,
XX kidney, bladder, breast, gastric, ovarian, colorectal, prostate,
XX pancreatic, lung, vulva, thyroid, hepatic carcinomas, sarcomas,
XX glioblastomas, and various head and neck tumours), leukaemias and
XX lymphoid malignancies, other disorders such as neuronal, glial,
XX astrocytal, hypothalamic and other glandular, macrophagal, epithelial,
XX stromal and blastocoelec disorders, and inflammatory, angiogenic and
XX immunologic disorders. AAC58242 to AAC58366 represent PCR primers and
XX hybridisation probes used in the isolation of the human PRO sequences.
XX AAC58367 to AAC58396 and AAB24057 to AAB24089 represent human PRO
XX polynucleotide and protein sequences given in the exemplification of the
XX present invention
XX
XX Sequence 1509 BP; 297 A; 451 C; 429 G; 332 T; 0 U; 0 Other;
XX

XX 07-JUN-1996; 96US-00660451.
XX 07-JUN-1995; 95US-00484722.
XX (MERI) MERCK & CO INC.
XX
XX Elliott KJ, Harpold MW;
XX WPI; 2003-511917/48.
XX P-PSDB; ADA10874.
XX
XX New nucleic acid encoding an alpha-6 or a beta-3 subunit of a human
XX neuronal nicotinic acetylcholine receptor (nAChR), useful for identifying
XX compounds that modulate human neuronal nAChR activity.
XX
XX Claim 33; Col 67-72; 63pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule comprising a
XX nucleotide sequence encoding an alpha 6 or a beta 3 subunit of a human
XX neuronal nicotinic acetylcholine receptor nAChR. nAChR's form ligand-
XX gated ion channels that mediate synaptic transmissions between nerve and
XX muscle and between neurons upon interaction with the neurotransmitter
XX acetylcholine. The nucleic acid molecule is useful for identifying
XX compounds that modulate human neuronal nAChR. The present sequence
XX represents DNA encoding the human neuronal nicotinic acetylcholine
XX receptor, nAChR, alpha 7 subunit. Note: the present sequence decodes to
XX the protein shown in ADA10874 not the one described in the specification
XX as being the nAChR alpha 7 subunit ADA10865.
XX
XX Sequence 1876 BP; 369 A; 553 C; 531 G; 423 T; 0 U; 0 Other;
XX
XX Query Match 24.9%; Score 374; DB 8; Length 1876;
XX Best Local Similarity 55.1%; Pred. No. 1.9e-93;
XX Matches 809; Conservative 0; Mismatches 635; Indels 24; Gaps 3;
XX
XX 33 GCGTTTCTGCTCGGATCGGAGCAAGTCTCTACGAGAGAGACTCTCTGAACGCGTTGCT 92
XX 114 GTCGCTCTGACGCTGTCCTGCAAGCGGAGTTCACAGGAGGCTTTTACAGGAGCTGGT 173
XX 93 GCGGAACTCAACACCTCGAGGACCGGTGGCCACGAGAGGCAACCGCTAGAGTTCAG 152
XX 174 CAGAACTCAATCCCTTGAGAGGCGCCGTGGCCAAATGACTCGCAACCACTCACCGTCTA 233
XX 153 GTTCGGCTTGACCTTGACGAAATCATTTGACGTGGACGAGAAATCAACTACTATAAC 212
XX 234 CTTCCTCTGAGCTCTCTGACATCATGAGCTGAGTGGAGAGCAACCAAGTTTAAACAC 293
XX 213 CAATATATGCTGTGTTGAGTGGAAATGACTCAACCTGAGTGGAAACGACAGCGAGTA 272
XX 294 CAACATTTGGCTGCAAAATGCTTTGGACAGATCACTATTACAGTGGAAATGTGCAGAA 353
XX 273 TGGCGGGTCAAGGACTCAGGATCAGCCCAACAGTTGTGGAAGCGGACGCTCTTAT 332
XX 354 TCAGGGGTGAACACTGTTGCTTCCAGATGGCCAGATTGGAACCAACACATTTCTCT 413
XX 333 GTATAATAGTGTGACGAGGTTTTCAGCGGACCTACAGACCAACAGTGGTGGTCAAG 392
XX 414 CTATAACAGTGTGATGAGCGCTTTGACGCGCACATCCACACTAACGTTGTGGTGAATC 473
XX 393 CGCGCGCAGTTCCTGTAGTGCACCTGGCCATTTCAAGAGCACATGCAAGATGACAT 452
XX 474 TTCTGGGCATTCGCAAGTACCTGCTCCAGGCATATTCAAGAGTTCCTGTACATCGATGT 533
XX 453 CGCGTGGTTTCCCTTCGACACCAACACTGTGATGTAAGTTTCGTAGCTGGACATATGA 512
XX 534 ACCTGGTTTCCCTTTGATGTGACGACTGCAACTGAAATTTGGTCTCTGGTCTTACGG 593
XX 513 CGGCAATCAGTTGGATCTGTGTGTTAAAGATAGGCGCGGCAATCTATCGGACTTCAT 572
XX 594 AGGCTGTCTCTGGATCTGACATGCGAGTA-----GGCAGATATCAGTGGCTATAT 644
XX 573 AACAAATGGGAGTGGTATCTAATAGATGCCAGGCAAAAGAACAAATAACATACGCG 632

Db 645 CCCAATGGAGATGGGACCTAGTGGGAATCCCGGCAAGAGGAGTGAAGGCTTCTATGA 704
Qy 633 GTGCTGCCCGAGCCCTAAGTGGAGCTCACTTCCACCATATGATAAAGAACGACCTT 692
Db 705 GTGCTGCAAGAGCGCTTACCCCGATGTCACTTACAGTACCACATGCGCCGACGCT 764
Qy 693 GTACTACTTCTCAACCTGATCGCTCGGTGGTGTGATCTCATCGATGGGACTCTCTCGG 752
Db 765 CTACTATGGCTCTCAACCTGCTGATCCCTGTGTGCTCATCTCCGCTCTGCGCTGTGCT 824
Qy 753 CTTTACACTGCCACAGACTCCGAGAGAACTCACACTTGGAGTCACTATTTCTTCTATC 812
Db 825 GTTCTCTCTTCTGACAGATTCGCGGAGAGATTTCCCTGGGATACAGCTTCTACTCTC 884
Qy 813 GCTGACGCTGTCTCAACCTGGTAGCCGAGACCTGTCACAGCTCTCCGACGCTATCCC 872
Db 885 TCTTACCGTCTTCTATGCTGCTGCTGATCATGCGCGCAACATCCGATTCGCTACC 944
Qy 873 CCTGTTAGGACGCTACTTCAATTTGATCATGTTTCACTGATGCTGCTGCTGCTGCTGCTG 932
Db 945 ATTGATAGCCAGTACTTTCGCGAGCACCATGATCATGCTGGGCTCTCGGTGGTGGTAC 1004
Qy 933 TGTGTTGGTACTCAATTTACCAACCATCGAACAGCTGATATATGATGATGATGATGATGAT 992
Db 1005 GGTGATCGTCTGACGATACCAACACGACCGCGGCGGCAAGATGCCAAGTGGAC 1064
Qy 993 AATATCAGTATTTCTTACAAATGGTTCATGATGATGCTGATGCTGATGCTGATGCTGATG 1052
Db 1065 CAGAGTATCTCTTCTGAACTGGTGGCGGTGTTCTCTGCAATGAAGAGGCGCGGGAGGA 1124
Qy 1053 GATCACCGAGAGACTATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1112
Db 1125 CAGGTGCGCGCGCGCTGCGCAGCAAGAGCGCGCTGCGAGCTGCGCAGTGTGGAGAT 1184
Qy 1113 GTGCTGCAAGTCTTCTGCTGCGGAATGTTCTAGATATTTGATGATGATGATGATGATGATGAT 1172
Db 1185 GAGCGCGTGGCG 1244
Qy 1173 TCGCGCTCTTAAAGTACTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1232
Db 1245 CCTGAGCGCGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1302
Qy 1233 CAGCGATTTTCGCTGCGGTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1292
Db 1303 GCCTGCTCCCG 1360
Qy 1293 GGTGAGCCACCTGCGGAGCTGCACTCTATCTATGAGAGAGCTGAGTTCATCAGCGCGCAG 1352
Db 1361 -----ACCGGACTTGGCCAAAGATCTCTGAGGAGGTCCGCTACATTTGCCAATCG 1409
Qy 1353 GATGAAGAGGCTGATGAGGAACCGGAGCTGATGAGGAGTTCGAGGAGTTCGCTGGGATGGT 1412
Db 1410 CTTCCGCTGCGAGGACGAAAGCGAGCGGCTCTGACGAGTGGAGTTCGCGCGCTGTGT 1469
Qy 1413 TGTGATAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1472
Db 1470 GGTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1529
Qy 1473 CTTGTTATCGGACCGCATATCATCTG 1500
Db 1530 CTTGATGCTGGCTCCCAACTTCTGTTGAG 1557

RESULT 10

AAV44687

ID AAV44687 standard; cDNA; 1590 BP.

XX

XX AAV44687;

XX AC

XX 09-OCT-1998 (first entry)

XX

XX V274T variant human alpha7 nAChR coding sequence.

Db 1348 TCGCTGCCAGGACGAAGCGAGCGGTCTGTCAGCGAGTGGAAGTTCGCGCGCTGTGG 1407
 QY 1415 TTGATAGTGTTCCTGCTGTTGTTACACTTTTACAAATCATCGGACAGTAGCTGCC 1474
 Db 1408 TGGACCGCTGTGCTCATGCGCTTCCTGCTTTTCCACCATCATCTGCACCATCGGCATCC 1467
 QY 1475 TGTATCGGCACGCGCATATCATGTG 1500
 Db 1468 TGAATCGGCTCCCAACTTGTGGAG 1493

RESULT 11

AAV12197
 ID AAV12197 standard; cDNA; 1876 BP.
 XX
 AC AAV12197;
 XX
 DT 14-MAY-1998 (first entry)
 XX
 DE Human neuronal nicotinic acetylcholine receptor alpha-7 subunit cDNA.
 XX
 KW Human; neuronal nicotinic acetylcholine receptor; alpha-7 subunit;
 KW brain tissue; screening; NACHR; antibody; ds.
 XX

OS Homo sapiens.

PH Key Location/Qualifiers
 FT CDS 73..1581

FT /*tag= a
 FT /product= "neuronal nicotinic acetylcholine receptor
 FT alpha-7 subunit"

XX WO9420617-A2.

XX 15-SEP-1994.

XX 08-MAR-1994; 94WO-US002447.

XX 08-MAR-1993; 93US-00028031.

XX (SALK) SALK INST BIOTECHNOLOGY IND ASSOC.

XX Elliott KJ, Ellis SB, Harpold MM;

XX WPI; 1994-303024/37.

XX P-PSDB; AAW44153.

PT Human neuronal nicotinic acetylcholine receptor subunits and DNA - also
 PT transformed cells useful for screening cpds. which modulate activity of
 PT the receptor.

XX Claim 8; Page 78-79; 99pp; English.

XX The present sequence encodes a human neuronal nicotinic acetylcholine
 CC receptor (NACHR) subunit. The cells expressing the alpha and/or beta
 CC NACHR subunits may be used in a method of screening compounds to identify
 CC any which modulate the activity of human neuronal NACHR. Subunit specific
 CC antibodies may be used to monitor the distribution and expression density
 CC of various subunits in normal vs diseased brain tissues. Testing of
 CC single receptor subunits or specific receptor subunit combinations with a
 CC variety of potential agonists or antagonists provides information with
 CC respect to the function and activity of the individual subunits and
 CC should lead to the identification and design of compounds that are
 CC capable of very specific interaction with one or more receptor subtypes.
 CC The resulting drugs should exhibit fewer unwanted side effects than drugs
 CC identified e.g. screening with cells that express a variety of subtypes

XX SQ Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;

Query Match

Best Local Similarity 55.08; Pred. No. 2.5e-93;

Matches 808; Conservative 1; Mismatches 635; Indels 24; Gaps 3;

QY 33 GGCTTTCTGCTCCGCTATCGGAGCAAGGTCTCTCAGAGAAGAGACTCTCTGACGCGTTGCT 92
 Db 114 GTGCTCTCTGCAAGCTGTCCCTGCAAGCGAGTTCAGAGAAGCTTTTACAGAGGAGCTGGT 173
 QY 93 GGGAACTCAACACACCCCTGGAGCGACCGGTGGCCCAACAGAGCGAACCCTAGAGGTGAG 152
 Db 174 CAAGAACTACAATCCCTTGGAGAGGCCCGTGGCCAAATGACTCGCAACCACTCACCGTCTA 233
 QY 153 GTTCGGCTTGAACCTTGCAGCAAAATCATTTGACGTGGACGAGAAGAAATCAACTTATTAAAC 212
 Db 234 CTTCTCCCTGAGCCTCTCTGAGATCATGACGTGGATGAGAAGAACCAAGTTTTAAACCAAC 293
 QY 213 CAATATATGCTGCTGTTGGAGTGAATGACTACAACCTGAGTGGACACACGAGTA 272
 Db 294 CAACATTTGGCTGCAAAATGCTTGGACAGATCACTATTTACAGTGAATGTGTGAGAATA 353
 QY 273 TGGCGGGTCAAGGACCTCAGGATCAGCCCCAACAAAGTTGTGGAAGCGGACGCTCTTAT 332
 Db 354 TCCAGGGGTGAAGACTGTTCTGTTTCCAGATGGCCAGATTTGGAACCAACAGACATCTTCT 413
 QY 333 GTATATATGCTGACGAGGTTTGTGACGGACCTTACAGACCAACCTGTTGTTGTCAGAAG 392
 Db 414 CTATAACAGTGTGATGAGCGCTTTGACGCCACATTCACACTAACCTGTTGTTGTAATTC 473
 QY 393 CGCGGCAGTTCCTGTACGTGCCCTGACCTGACATATTCAGAGACACATGCAAGATGGACAT 452
 Db 474 TTCTGGGCATTTGCCAGTACTGCTCTCAGGCATATTCAGAGTTCCTGCTACATCGATGT 533
 QY 453 CGCGTGTTCCTTTCGACGACCAACACTGTGATATGAAGTTGCGTGTGAGTGGACATATGA 512
 Db 534 ACGCTGTTTCTCCCTTTGATGTGCAGCACTGCAAACTGAAAGTTTGGTCTGCTTACGG 593
 QY 513 CGGCAATCAGTTGGATCTGCTGTCTAAAGATGAGGAGCGGCGGATCTATCGGACTTCAT 572
 Db 594 AGGCTGTCTCTGTGATCTGCAGATGCAGGA-----GCGAGATATCAGTGGCTATAT 644
 QY 573 AACAAATGGGGAGTGGTATCTAATAGGAATGCCAGGCAAAAAGAAACAATAACATACGC 632
 Db 645 CCCNATGGAGATGGACCTAGTGGAAATCCCGGCAAGAGGATGAAGTTCATATGA 704
 QY 633 GTGCTGCCCGAGCCCTTACGTGGACGTCACCTTTCACCATCATGATGAAGAAGACGAACTT 692
 Db 705 GTGCTGCAAGAGAGCCCTTACCCGATGTCACTTTCACAGTCACTATGCGCGCAGACGCT 764
 QY 693 GTACTACTTCTTCAACCTGATCTCCGTCGGTGTCTGATCTCATCGATGCGACTCTCTCGG 752
 Db 765 CTACTATGGCTCAACCTGCTGATCTCCCTGTGCTCATCTCCGCTCTCCCTGCTGGT 824
 QY 753 CTTTCACACTGCCACCACTCCCGAGAGAACTCACACTTTGGAGTCACTATTCTTCTATC 812
 Db 825 GTTCCTGCTCTTCGAGATTCGGGGAGAGATTTCCCTGGGAGATAACAGTCTTACTCTC 884
 QY 813 GCTGAGCGTGTCTCTCAACCTGTTAGCCGAGACCTGCGCACAGGTCTCCGACGCTATCCC 872
 Db 885 TCTTACCGTCTTTCATGCTGCTGCTGGTGGATCATGCGCGCAACATCCGATTCGGTACC 944
 QY 873 CCTGTTAGGACCTACTTCAATTGCATCATGTTTCATGTGTCGTCGCTCTGTTGTTACTGAC 932
 Db 945 ATTGATAGCCAGTACTTCCCGCAGCCATGATCATGCTGGGCTCTTCGGTGGTGGTAC 1004
 QY 933 TGTGGTGTACTCAATTTACCCCATCGAACAGTGTATATACATGAATGCCACAGTGGAT 992
 Db 1005 GGTGATCGTCTGCTGAGTACCAACCAACCGACCCCGGCGGAGAGATGCCAAGTGGAC 1064
 QY 993 AAAATCAGTATTCCTACATGTTGCGATGATGCTGCGAATGTGAGGCGCAGGAGAA 1052
 Db 1065 CAGAGTCACTCTTCTGAACCTGGTGGCGGTGTTCTCTSCGAATGAAGAGGCCCGGGGAGGA 1124
 QY 1053 GATCACCGAGGAGACTATATATGATGAACACGAGATGAGGAGCTGGAACCTGAAGGAGAG 1112
 Db 1125 CAAGGTGCGCCCGCTGCCAGCACAGAGCGGCGCTGCAGCTGCGCAGTGTGGAGAT 1184

QY 1113 GTCTGGAAGTCTTCTGGCGAATGTTCTAGATATTGATGATGACTTTCAGACACGGCCC 1172
Db 1185 GAGCGCCGTGGCGCCGCCGCCAGCAACCGGAACTGCTGTATATCGGCTTCGCGG 1244
QY 1173 TCGCCTCTTACAGTACTGCTTCGACCGGGAATTTGGGACCTGGGTGCTCAATATTCG 1232
Db 1245 CTTGGACGGGTGCACTGTGTCGCCGCCCGGACTCTGGGGTAGTGTGTGGCCGAT--G 1302
QY 1233 CAGGATTTCCGTGCGGTGCTGCTCGTCCGTCAGATGGAAGAGTGGGCGCGGGCT 1292
Db 1303 GCCTGCTCCGCCACGACGATGACCTCTCTCAGCGCGGCAACCCCGGAGGGG-- 1360
QY 1293 GGGTAGCCACCTCGGAGCTGCACCTCATCTACTGAGAGAGCTGCAGTTTCATCAGGCCAG 1352
Db 1361 -----ACCGGACTTGGCCAGATTCCTGGAGAGGTCCGCTACATTCGCAATCG 1409
QY 1353 GATGAAGAAGGCTGATGAGGAAGCCGAGCTGATCAGCAGCTGGAAGTTTCTCGGATGGT 1412
Db 1410 CTTCCGCTGCCAGACGAAGAGGCGGTCTGACGGAGTGAAGTTCCGCCCTGTGT 1469
QY 1413 TGTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1472
Db 1470 GGTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1529
QY 1473 CTTGTTATCGGACCGCATATCATCTG 1500
Db 1530 CTTGATGCTGGCTCCCACTTCTGGAG 1557

RESULT 12

ID ABS54875
ID ABS54875 standard; cDNA; 1876 BP.

XX AC ABS54875;

XX 06-DEC-2002 (first entry)

DE Human neuronal nicotinic acetylcholine receptor alpha 7 subunit cDNA.

XX Human; neuronal nicotinic acetylcholine receptor; nNACHr; gene; ss;

XX ion flux; alpha 7 subunit.

XX Homo sapiens.

XX Key Location/Qualifiers
FH CDS 73..1581

FT /*tag= a

FT /product= "Human nNACHr alpha 7 subunit"

XX US6440681-B1.

XX 27-AUG-2002.

XX 07-JUN-1995; 95US-00487596.

XX 03-APR-1990; 90US-00504455.

XX 30-NOV-1992; 92US-00938154.

XX 08-MAR-1993; 93US-00028031.

XX 08-NOV-1993; 93US-00149503.

XX (MERI) MERCK & CO INC.

XX Elliott KJ, Ellis SB, Harpold MM;

XX WPI; 2002-711528/77.

XX P-PSDB; ABG70492.

XX Identifying antagonists or agonists of human neuronal nicotinic

PT acetylcholine receptors, by contacting recombinant cells with test

PT compound, and measuring ion flux of cells or binding of compound to

PT nNACHr.

XX Claim 101; Col 57-60; 56pp; English.

XX The invention relates to a method for identifying compounds that are
CC antagonists or agonists of human neuronal nicotinic acetylcholine
CC receptors (nNACHrs), by contacting recombinant cells with a test compound
CC and measuring ion flux, the electrophysiological response of the cells or
CC binding of the test compound to the nNACHr. The recombinant cells are
CC produced by transfection with a nucleic acid encoding at least one human
CC nNACHr (alpha or beta) subunit, such that the cells express an nNACHr
CC comprising one human subunit encoded by the transfected nucleic acid.
CC This sequence represents cDNA encoding the alpha 7 subunit of the human
CC nNACHr polypeptide

XX Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;

Query Match 24.9%; Score 373.6; DB 6; Length 1876;

Best Local Similarity 55.0%; Pred. No. 2.5e-93;

Matches 808; Conservative 1; Mismatches 635; Indels 24; Gaps 3;

QY 33 GGCTTTGCTGCCGTATCGGACCAAGTCTCTCAGAGAGAGACTCTTGACCGTTGCT 92

Db 114 GTCGCTCCTGACGTGCTCCCTGCAAGGCGAGTTCAGAGGAAGCTTTACAGGAGCTGGT 173

QY 93 GCGCAACTACAAACACCCCTGGAGCGACCGGTGGCCAAACGAGAGCGAACCGCTAGAGGTCAG 152

Db 174 CAAGAATACATCCCTTGGAGAGGCGCGTGGCCATGACTCCGAACCACTCCACCGTCTA 233

QY 153 GTTCGGCTTGACCTTGGAGCAAAATCATTTGACGTGGACGAGAAATCAACTATTATPAC 212

Db 234 CTTCTCCCTGAGCCTCTCGAGATCATGGACGTGGATGAGAGAACCAAGTTTAAACCAC 293

QY 213 CAATATATGGCTGCTCGTTGGAGTGAATGACTACACCTGAGGTGGAACGACAGCGAGTA 272

Db 294 CAACATTTGGCTGCAAAATGTCTTGGACAGATCACTATTTACAGTGAATGTGTFCAGAA 353

QY 273 TGGCGGGGTCAAGGACCTCAGATCACGCCCAAGATTGTGGAAGCGGACGCTCCTTAT 332

Db 354 TCCAGGGGTGAGACTGTTCTGTTCCAGATGCCAGATTTGAAACACAGACATTTCTTCT 413

QY 333 GTATAATAGTGTGACGAGGGTTTTCACGGGACCTACAGACCAACGTTGGTGTGTGAGAAG 392

Db 414 CTATAACAGTGTCTGATGAGCGCTTTGACGCCACATTCACACATTAAGCTGTGTGTAATTC 473

QY 393 CGCGGCGAGTGGCTGTGCTGACCTGCGCATATTCAGAGCACATGCAAGATGGACAT 452

Db 474 TTTCTGGGCAATTCGCACTGCTGCTCCAGCATATTCAGAGTTCCTGCTACATCATGT 533

QY 453 CGCGTGGTTTCCCTTCGACGACCAACACTGTGATATGAAGTTTCGTAGCTGACATATGA 512

Db 534 ACGCTGGTTTCCCTTTGATGTGCGACACTGCAAACTGAAGTTTGGGTCTGTGCTTACGG 593

QY 513 CGGCAATCAGTTGGATCTGGTCTGCTAAAGATGAGGCGAGCGCGGCGATCTATCGGACTTCAT 572

Db 594 AGGCTGGTCTCTGGATCTGCAGATGCAGGA-----GGCAGATATCAGTGGCTATAT 644

QY 573 AACAAATGGGAGTGTCTATCTAATAGGAATGCCAGGCAAAAGAACACATATACATACGC 632

Db 645 CCCCAATGAGGAATGGGACCTAGTGGAAATCCCCGGGAAGAGAGTGAAGGTTCTATGA 704

QY 633 GTGCTGCCCCGAGCCCTACGTGACGCTCACCTTCACCATCATGATGAAGAGAGCAACCTT 692

Db 705 GTGCTGCAAGAGCCCTACCCCGATGTCACTTTCAGTGCACCATGCGCGCAGGAGGCT 764

QY 693 GTRACTACTTCTCAACTGATGCTCCGTCGCGTGTGATCTCATCTGAGTCACTATTCTTCTATC 752

Db 765 CTACTATGGCCTCAACCTGCTGATCCCTGTGTGCTCATCTCCGCCCTCGCCCTGTGCT 824

QY 753 CTTTCACTGCCACAGACTCCGGAGAGAACTCACACTTGGAGTCACTATTCTTCTATC 812

Db 825 GTTCTGCTTCTGCGAGATTCGGGGAGAGATTTCCCTCGGGGATACAGTCTTACTCTC 884

QY 813 GCTGACGGTGTCTCCTCAACTGGTAGCGAGACCTCTGCCACAGGTCTCCGACGCTATCCC 872

Db 885 TCTTACCGTCTTCATGCTGCTGTGCTGAGATCATGCCGCAACATCCGATTCGTTACC 944

645 CCCAATGAGATGGACCTAGTGGGAATCCCGCAAGAGAGTGAAGGTTCTATGA 704
QY
633 GTGCTCCCGAGCCCTAGTGACGCTACCTTCAACATCATATAGAAAGACCACTT 692
Db
705 GTGCTGCAAGAGCCCTACCCGATGTCACTTACAGTGACCATGCGCGCAGGACGCT 764
QY
693 GTACTACTTCTTCAACTGATCGTCCGCTGCGTGTGATCTCATGATGCACTCTCCG 752
Db
765 CTACTATGGCTCAACTGCTGATCTGTGTGCTCATCTCCGCTCGCCCTGTCTGGT 824
QY
753 CTTTACACTGCCACGAGCTCCGGAGAGAACTCACACTTGGAGTCACTATTCTTCTATC 812
Db
825 GTTCTGCTTCTCGAGATTCCGGGAGAGATTTCCTTGGGATTAACAGTCTTACTCIC 884
QY
813 GCTGACGGTGTCTCTCAACTGCTAGCGGAGACCTTGCACACAGGTCTCCGACGCTATCCC 872
Db
885 TCTTACCGTCTTCACTGCTGCTGCTGAGATCATGCCCGCAACATCCGATTCCGTACC 944
QY
873 CTTGTTAGGAGCTTCAATTGCAATGCTATGTTTCAATGCTAGCTGCTGTGCTACTGAC 932
Db
945 ATTGATAGCCAGTACTTCCGACACCATGATCATCTGTTGGGCTCTCGTGTGGTGAC 1004
QY
933 TGTGTTGGTACTCAATTACCACCATCGAACAGCTGATATACATGAATGCCACAGTGGAT 992
Db
1005 GGTGATCGTCTGACGTACCAACACCAACGACCGCGGGCAAGATGCCAAGTGGAC 1064
QY
993 AATAACAGTATTCTCAATGTTGCCATGGATGACTGCGAATGTCGAGGCCAGGGAAGAA 1052
Db
1065 CAGAGTCATCTTCTGAACTGTGCGGTGTCTTCCSCGAATGAAGAGGCCCGGGAGGA 1124
QY
1053 GATCACCAGGAAGACTAATGATGAACAGAGAGATGAGGAGCTGGAATGAAGAGAG 1112
Db
1125 CAAAGTGGCCCGCTGTCAGCAGACAGACGCGCGCTGAGCTGCGCAAGTGTGAGAT 1184
QY
1113 GTCGTCGAAGTCTTGTGTCGCAATGTTCTAGATATTGATGACTTCAGACACGCGCC 1172
Db
1185 GAGCGCGTGGCGCGCGCGCGCGCAACGGAACCTGCTGATCATCGCTTCCGCG 1244
QY
1173 TCGCCTCTTAACAGTACTGCTCGACCGGGAATTTGGGACCTGGGTGCTCAATATTCGG 1232
Db
1245 CTTGACGGCGTGCACGTGTCTCCGACCCCGCACTCTGGGGTAGTGTGTGGCGCAT--G 1302
QY
1233 CAGGATTTTCGTCGCTGCTCGTCCGTCAGTGGAGAGAGTGGCGCGCGGCT 1292
Db
1303 GCTGCTCCCGACGACGATGAGACCTCTCGACGCGGGGCAACCCCGGAGGGGG-- 1360
QY
1293 GGGTAGCCACATCGGAGCTGCACCTCATACTAGAGAGCTGCAGTTTCATACGCGCAG 1352
Db
1361 -----ACCGGACTTGGCCAAGATCCTGGAGAGGTCCGCTACATTGCCAATCG 1409
QY
1353 GATGAAGAGGCTGATGAGGAACCGAGCTGATCAGGAGCTGGAAGTTTCTCGGATGGT 1412
Db
1410 CTTTCCGCTGCGAGGACGAAGAGGCGGCTGACGCGAGTGAAGGTTTCCCGCCTGTGT 1469
QY
1413 TGTGTGATAGTTTGTGCTGTGTGTTTACACTTTTACAAATCATCGCAGCAGTAGTGT 1472
Db
1470 GGTGAGCGGCTGTGCTCATGGCTTCTCGTCTTCCCATCATCTGACCAATCGGCAT 1529
QY
1473 CTTGTTATCGGACCGCATATCATCGTG 1500
Db
1530 CTTGATGTGGCTCCCAACTTCGTGGAG 1557

RESULT 14

AAC90385

ID AAC90385 standard; cdna; 1509 BP.

XX

AC AAC90385;

XX

DT 14-MAR-2001 (first entry)

XX

DE Mutant human alpha7 ligand gated ion channel coding sequence #1.

XX

KW Human; alpha7 nicotinic acetylcholine gated ion channel; mutant;
KW 5-hydroxytryptamine; 5-HT3; calcium ion conductance; ss.
XX Homo sapiens.
OS
PN WO200073431-A2.
XX
PD 07-DEC-2000.
XX
PF 25-MAY-2000; 2000WO-US011862.
XX
PR 27-MAY-1999; 99US-0136174P.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
XX Groppi VB, Wolfe ML, Berkenpas MB;
PI P-PSDB; AAB50015.
XX
DR WPI; 2001-061524/07.
DR P-PSDB; AAB50015.
XX

Special cell culture medium for treating cells and for inducing mammalian cell lines to conduct calcium ions, comprising specified concentrations of ions of sodium, calcium and potassium at specified pH.

Claim 50; Page 69; 77pp; English.

The present sequence is the coding sequence for a mutant human alpha7 nicotinic acetylcholine gated ion channel. The human alpha7 ion channel was used in the construction of an alpha7/5-hydroxytryptamine (5-HT3) chimeric ligand gated ion channel (see AAC90382 and AAB50014). The alpha7/5-HT3 chimeric ion channel can be expressed by recombinant cells in the present invention, resulting in preferential calcium ion conductance by the cells. The protein encoded by this sequence has the wild-type threonine residue at position 230 substituted by a proline residue

Sequence 1509 BP; 296 A; 452 C; 429 G; 332 T; 0 U; 0 Other;

Query Match 24.8%; Score 372.4; DB 4; Length 1509;
Best Local Similarity 55.0%; Pred. No. 4.9e-93;
Matches 808; Conservative 0; Mismatches 636; Indels 24; Gaps 3;

QY 33 GGCTTTGTCGCCGTATCGAGCAAGTCTCTCAGAGAGAGACTCTCTGACCGGTGCT 92
Db 42 GTCGCTCTCGACGTGTCCCTGCAAGCGAGTTCCAGAGGAAGCTTTACAGGAGCTGT 101
QY 93 GCGCAACTACAACACCTCGAGCGACCGGTGGCCACGAGAGCGAACCGCTAGAGTCTAG 152
Db 102 CAAGAACTACAATCCCTTGAGAGAGGCCCGTGGCCATGACTCGCAACCACTCACCGTCTA 161
QY 153 GTTCGGCTTGACCTTTCAGCAAAATCATTTGACGTGGACGAGAGAAATCAACTACTTATAAC 212
Db 162 CTTCTCCCTGAGCTCTCTGCAGATCATGACCGTGGATGAGAGAACCAAGTTTAAACCAC 221
QY 213 CAATATATGGCTGCTCGTTGGAGTGGAAATGACTACAACTGAGGTGAGAACGACGAGTA 272
Db 222 CAACATTTGGCTGCAAAATGTTTGGACAGATCACTATTTACAGTGAATGTGTCAANA 281
QY 273 TGGCGGGGTCAAGGACCTCAGGATCACGCCCAACAAAGTTGTGGAAGCCGAGACCTCTTAT 332
Db 282 TCCAGGGGTGAGACTGTTTCGTTTCCAGATGCCAGATTGGAAACCAAGACATCTTCT 341
QY 333 GTATAATAGTGTGACGAGGGTTTTCAGCGGACCTTACAGCAACCAAGCTGGTGTGTAAG 392
Db 342 CTATAACAGTGTGATGAGCGCTTTTGACGCCACATTCACACTTAACGTGTGTGTAATTC 401
QY 393 CGGCGGAGTTGCTGTAGTCCACTGTCATATTTCAAGAGCACATGCAAGTGTGACAT 452
Db 402 TTCTGGGCATTTGCCAGTACTCGCTCCAGGCATATTTCAAGAGTTCCTGTACATCATGT 461
QY 453 CGGCTGTTTCCCTTCGACGACCAACACTGTGATATGAAGTTTCGGTAGCTGGACATATGA 512
Db 462 ACCTGTTTCCCTTTGATGTGAGCACTGCAAACTGAAGTTTGGTCTGCTTACCG 521

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 14:08:32 ; Search time 4039.19 Seconds

(without alignments)

11111.850 Million cell updates/sec

Title: US-09-303-232-5_COPY_95_1597

Perfect score: 1503

Sequence: 1 atggccctatgttgcggc.....caccgcataatcatgtgcaa 1503

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lasting first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_estl:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gssl:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	423.8	28.2	885	12	BG632919
2	331.2	22.0	607	9	AI292581
3	302.2	20.1	2296	14	CD013901
4	301.2	20.0	1201	9	AL530299

5	294.8	19.6	1436	29	AY402873	Homo sapi
6	293.2	19.5	2940	11	AK034228	Mus muscu
7	293.2	19.5	3230	11	AK083157	Mus muscu
8	286.8	19.1	1436	29	AY402874	Pan trogl
9	281.6	18.7	3483	11	AK081254	Mus muscu
10	281.6	18.7	4037	11	AK049722	Mus muscu
11	281.6	18.7	4046	11	AK051742	Mus muscu
12	275.6	18.3	1864	11	AK053497	Mus muscu
13	275.6	18.3	2916	11	AK051730	Mus muscu
14	275.6	18.3	3126	11	AK080415	Mus muscu
15	274.4	18.3	1442	29	AY402876	Homo sapi
16	272.4	18.1	1436	29	AY402875	Mus muscu
17	267	17.8	1442	29	AY402877	Pan trogl
18	258.8	17.2	2010	11	AK080475	Mus muscu
19	256.6	17.1	833	14	CB245337	UI-M-FY0-
20	253	16.8	1454	29	AY402878	Mus muscu
21	246.2	16.4	1374	29	AY406232	Mus muscu
22	246.2	16.4	4290	11	AK029177	Mus muscu
23	240	16.0	908	13	BUI49265	AGENCOURT
24	239.8	16.0	1374	29	AY406230	Homo sapi
25	239.2	15.9	922	13	BUI15857	AGENCOURT
26	239.2	15.9	1466	14	CD013891	90139069
27	238.2	15.8	1603	29	AY411327	Homo sapi
28	236.8	15.8	1036	13	BX437801	BX437801
29	234.8	15.6	755	13	BUI702422	UI-M-FY0-
30	231.6	15.4	902	29	AY407186	Mus muscu
31	229.4	15.3	759	14	CD804155	UI-M-HB0-
32	229	15.2	730	14	CF742344	UI-M-HB0-
33	228.2	15.2	615	14	CB149460	K-EST0205
34	228.2	15.2	1374	29	AY406231	Pan trogl
35	220.4	14.7	2513	11	AK033068	Mus muscu
36	219	14.6	1603	29	AY411329	Mus muscu
37	216.6	14.4	902	29	AY407184	Homo sapi
38	216.2	14.4	819	14	CF530720	UI-M-FY0-
39	212.6	14.1	1034	13	BX403124	BX403124
40	210.8	14.0	755	10	AW914206	EST345510
41	210.4	14.0	658	12	BM711715	UI-E-CL1-
42	210.4	14.0	797	14	CA326954	UI-M-FY0-
43	208.8	13.9	777	14	CF538185	UI-M-G10-
44	205.8	13.7	791	14	CF739825	UI-M-HD0-
45	202.2	13.5	939	13	BQ720344	AGENCOURT

ALIGNMENTS

RESULT 1	BG632919	885 bp	mRNA	linear	EST 23-APR-2001
LOCUS	GH16126.3	prime GH Drosophila melanogaster head	POT2	Drosophila	
DEFINITION	melanogaster cDNA clone GH16126 3 similar to CG4128: FBan0004128				
	'ion channel' located on: 2L 30D1-30E1; 04/10/2001, mRNA sequence.				
ACCESSION	BG632919				
VERSION	BG632919.1	GI:13758409			
KEYWORDS	EST.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 885)				
AUTHORS	Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G., Lewis,S. and Rubin,G.M.				
TITLE	BDGP/HMI Drosophila EST Project				
JOURNAL	Unpublished (2001)				
COMMENT	Other ESTs: GH16126.5prime				
	Contact: Stapleton, M.				
	BDGP				
	Lawrence Berkeley National Lab				
	One Cyclotron Rd, Berkeley, CA 94720, USA				
	Fax: 510 486 6798				
	Email: http://www.fruitfly.org/EST_estfruitfly.berkeley.edu				
	Based upon the presence of a XhoI site followed by a run of 14 or				

more T residues at the beginning of the sequence, this clone was polyadenylated. The resulting Poly-T sequence has been removed. hit genomic AE003511: arm:X [18792641..19136447]
estimated-cyto:18A3-18C6: 04/10/2001
Plate: GH.161 row: C column: 2
High quality sequence stop: 784.

FEATURES

Location/Qualifiers

source

1..885
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="GH16126"
/sex="male and female"
/dev_stage="adult"
/lab_host="DH5 - alpha"
/clone_lib="GH Drosophila melanogaster head pot2"

/notes="Organ: head; Vector: pot2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into pot2. Plasmid cDNA library."

ORIGIN

Query Match 28.2%; Score 423.8; DB 12; Length 885;
Best Local Similarity 68.9%; Pred. No. 1.7e-95;
Matches 581; Conservative 0; Mismatches 262; Indels 0; Gaps 0;

QY 161 TGACCTTGACGAAATCATTGACGTGACGAGAGAAATCAACTACTTATAACCAATATAT 220
DB 884 TAACTCATGACGATTATCGATGTGCACGAGAAATCAACTGCTTATACGAATATT 825
QY 221 GGCTGCTGTTGGAGTGGAAATGACTACAACTGAGGTGGACGACGAGTATGGGGGG 280
DB 824 GGCTCAATTTGGAATGGAACGATATGAATCTTCGATGGAATTCGAGTGAGTTCGGTGGTG 765
QY 281 TCAAGGACCTCAGATCAGCCCAACAAAGTTGTGGAAGCCGAGCGTCTTATGTATAAT 340
DB 764 TCGGGATCTCGAAATTCGCCACATCGCCTATGGAACCGGATGATGATGACAACA 705
QY 341 GTGCTGACGAGGCTTTGACGGGACCTTACCAGACCAACGTTGGTTCAGAACGGGGCA 400
DB 704 GTCCGACGAGGGCTTCGATGGAACGTACGCCCAATATGGTGGTTCGCAATATGGGA 645
QY 401 GTTGCCTGTACGTGCCACCTGGCATATTCAGAGCACATCAAGATGGACATCGCGTGT 460
DB 644 GCTGTCTGACGTACCGCCAGGTATATTTAAGTCAACGTGTAAGATCGACATACGTGT 585
QY 461 TTCCCTTCGACGACCACTGTGATATGAATTCGGTAGCTGACATATGACGGCAATC 520
DB 584 TTCCATTCGACGATCAGATGTGAATGAAATTTGGTTCGTGACCTACGATGGGTTTC 525
QY 521 AGTTGGATCTGGTGTAAAGATGAGCGCGCGCGATCTATCGGACTTCATTAACAAATG 580
DB 524 AGTTGACCTGAGTTGCAGGACGAGCTGGTGGCGACATTTCTAGCTTTATACCAATG 465
QY 581 GGGAGTGGTATCTAATAGGAATCCAGGCAAAAAGAACACAATAACATACGCGTGTGCC 640
DB 464 GCGAATGGGACTTGTAGGTGTGCCCGTAAACGAAATGAAATCTACTATAATGCTGCC 405
QY 641 CCGAGCCTTACGTGGAGCTTCACTTACCATATGATGAAGAAGACGACCTTGTACTACT 700
DB 404 CAGAACTTATATGACATACCAATTCGCCATTTGATAAGCGCAAAACGTTGTACTATT 345
QY 701 TCTTCAACTGATCTGTCGCGTGTGATCTCATCGATGGGACTCTCTCGGCTTCACAC 760
DB 344 TTTTCAATCTGATGTGCGGTGCTGATGATCGCTCCATGACACTGCTAGGTTTACAC 285
QY 761 TGGCACAGACTCGGAGAGAACTCACATTGGAGTCACTATTCTTCTATCTCGTGAACG 820
DB 284 TGGCACAGATTTCTGGTGAAGGCTTTTCGCTTGGAGTTACAATTTCTATTATCTCGTTACG 225
QY 821 TGTTCCTCAACCTGGTACGAGACCTGCGCACAGGTCTCGGAGCTATCCCTCTGTAG 880
DB 224 TCTTCCTCAACATGGTGGCGAACAATGCGGGACCTCCGATGCGGTACCGTGTCTCG 165

QY 881 GGACGTACTTCAATTGCATCATCTTTCATGTCAGTCGCTGTGTGTACTGACTGTGTGGTGG 940
DB 164 GAACATTATTCAATTGCATTAATGTTATGTTGTCGCTCATCAGTTGTGTCAACCATCTTG 105
QY 941 TACTCAATTACCAACCATCAAGAGCTGATATCATGAAATGCCACAGTGGATAAAATCAG 1000
DB 104 TCTCAATTATCATCATAGAAATCCAGATACGATAAGAAATGAGTGAATGGTAAATAGTTT 45
QY 1001 TAT 1003
DB 44 TAT 42

RESULT 2

AL292581

LOCUS

DEFINITION

GH15518.5prime GH Drosophila melanogaster head pot2 Drosophila melanogaster cDNA clone GH15518 5 similar to CG4128: F8an0004128, 'ion channel', located on: 2L 30D1-30E1; 04/10/2001, mRNA sequence.

ACCESSION

AL292581

VERSION

AL292581.1

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Other ESTs: GH15518.3prime

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST. est@fruitfly.berkeley.edu>

hit genomic AE003626: arm:2L [9617316..9882551]

estimated-cyto:30C7-30F4: 04/10/2001

Plate: GH.155 row: B column: 6

High quality sequence stop: 521

POLYA=No.

Location/Qualifiers

source

1..607

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/clone="GH15518"

/sex="male and female"

/dev_stage="adult"

/lab_host="DH5 - alpha"

/clone_lib="GH Drosophila melanogaster head pot2"

/notes="Organ: head; Vector: pot2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into pot2. Plasmid cDNA library."

ORIGIN

Query Match 22.0%; Score 331.2; DB 9; Length 607;
Best Local Similarity 76.7%; Pred. No. 3.1e-72;
Matches 405; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 11 TGTTCGCGCTTGGCGCTGCTGGCTTTGCTGCCGCTATCGGAGAGGTCTCTCACAGA 70
DB 75 TGTTCGCTGTGATCTTCTGCGGATATTAAGAAAGCTGTCAAGGACCTCAAGAA 134
QY 71 AGAGACTCTTGACGCGCTTCTCGGAACTACAAACCCCTGGAGCGACCGGTGGCCAAAC 130
DB 135 AGCGCTCTGTAACCATCTGCTCCACCTACAAATACGTGAGCGACCGGTGGCCAAATG 194
QY 131 AGAGCGACCGCTAGAGGTGAGTTCGGCTTGGCTTGCACCTTGCAGCAAAATCAITGACGTGACG 190

Db 195 AATCGGAGCCCTGGAGGTTAAGTTTCGAGTGAAGCTGACGCTGACGAGATCATCGAGCTGGAGC 254

QY 191 AGAAGAACTCACTATTATAACAATATATAGTGTGCTGTGTTGAGTGAAGTCACTCAACC 250

Db 255 AGAAGAACTCAGCTTCTCATACGAATCTTTGGCTTTCTGTTGAGTGAAGCACTACAATC 314

QY 251 TCAGGTGGAACACAGCAGTATGCGGGTCAAGGACCTCAGGATCAGCCCAACAAGT 310

Db 315 TCGCTGGATGAACGGAATACGCGGGTCAAGGATCAGGATCAGCCCAACAAGC 374

QY 311 TGTGAAGCCGACGCTCTTATATATATAGTGTGCTGACGAGGTTTGAAGGACCTACC 370

Db 375 TGTGAAGCCGACGCTCTTATATATAGTGTGCTGACGAGGTTTGAAGGACCTACC 434

QY 371 AGACCAACGTGTGTGTCAGAACGGCGGAGTGTGCTGACGAGGTTTGAAGGACCTACC 430

Db 435 ACACCAAGCTTGTGTGTCAGAACGGCGGAGTGTGCTGACGAGGTTTGAAGGACCTACC 494

QY 431 AGACCAACGTGTGTGTCAGAACGGCGGAGTGTGCTGACGAGGTTTGAAGGACCTACC 490

Db 495 AGACCAACGTGTGTGTCAGAACGGCGGAGTGTGCTGACGAGGTTTGAAGGACCTACC 554

QY 491 AGTTCGGTGTGTCAGAACGGCGGAGTGTGCTGACGAGGTTTGAAGGACCTACC 538

Db 555 AATTCGGTGTGTCAGAACGGCGGAGTGTGCTGACGAGGTTTGAAGGACCTACC 602

RESULT 3

CD013901

LOCUS 2296 bp mRNA linear EST 21-OCT-2003

DEFINITION 90134548 single gene library Homo sapiens cDNA, mRNA sequence.

ACCESSION CD013901

VERSION CD013901.1 GI:37777431

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1. (bases 1 to 2296)

Jin, P., Fu, G.K., Wilson, A.D., Yang, J., Chien, D., Hawkins, P.R., Au-Yang, J. and Stuve, L.L.

PCR isolation and cloning of novel splice variant mRNAs from known drug target genes

Unpublished (2003)

Contact: Jin, P.

Incyte Corporation

3160 Porter Drive, Palo Alto, CA 94304, USA

Tel: 650 621 8639

Fax: 650 621 8965

Email: pjin@incyte.com.

Location/Qualifiers

1..2296

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone_lib="Single gene library"

/note="Vector: pDrive Cloning Vector; RT-PCR was performed using gene-specific primers flanking the open-reading frame. PCR products were subcloned into pDrive Cloning Vector and sequenced completely using M13 forward and reverse primers. Sequencing gaps were closed by re-sequencing using primers flanking the gapped areas."

FEATURES

source

Query Match 20.1%; Score 302.2; DB 14; Length 2296;

Best Local Similarity 56.9%; Pred. No. 1.2e-64;

Matches 575; Conservative 0; Mismatches 433; Indels 3; Gaps 1;

QY 33 GCGTTTGTGCGCCGATTCGAGCAGAGTCTCTCAGGAAAGAGACTCTCTGAACGCGTGTCT 92

Db 316 GGCATTTCGCGAGGAGGCTCGCATACCGAGACTGAGGACCGGCTCTTCAAAACACCTCTT 375

QY 93 GCGGAACCTACACACCTCGAGCGCGGTGGCCCAAGAGAGCAACCGCTAGAGGTGAC 152

ORIGIN

RESULT 4

AL530299

LOCUS 1201 bp mRNA linear EST 23-MAY-2003

DEFINITION AL530299 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens cDNA clone CS0DD007YP05 5-PRIME, mRNA sequence.

ACCESSION AL530299

VERSION AL530299.2 GI:31068132

KEYWORDS EST.

SOURCE Homo sapiens (human)

Db 376 CCGGGCTACAAACCCCTGGCGCGCCGTCGCCCAACTTCAGAGCTGGTGAATGTGCG 435

QY 153 GTTCGGCTTGACCTTTGCAGCAAAATCAATTCAGCTGCAGCAAGAAATCACTACTATTAAAC 212

Db 436 CTTTGGACTGTCCATCGCTCAGCTCATCGATGTGAGTGAAGAAACCAATGATGACCAAC 495

QY 213 CAATATATAGTGTGCTGTTGAGTGAATGACTACAACTGAGTGAAGCAAGCAGGAGTA 272

Db 496 CAACCTCTGGCTTAAACACAGGAGTGGAGGACTTACAACTGCGCTGGAAACCCCGCTGATTT 555

QY 273 TGGCGGGTCAAGGACCTCAGGATCAGCCCAACAAAGTTGTGAAACCGGAGCTCTTAT 332

Db 556 TGGCAACATCATCTCTCAGGCTCCCTTCTGAGATGATCTGGATCCCGCAATTTGTTCT 615

QY 333 GTATAATAGTGTGACGAGGGTTTACGGGACCTTACAGCAACCAAGTGGTGGTCAAG 392

Db 616 CTACAACAATGCAGATGGGAGTTTTCAGTGAACCCACATGACCAAGGCCACCTCTTCTC 675

QY 393 CGGCGGAGTGTGCTGATGTCGACCTGGGATATTCAGAGCACATGCAAGATGGACAT 452

Db 676 CACGGGCACTGTGCACTGGTGTGCGCGGCCATCTTACAAGAGCTCTCTGAGCATCGAGCT 735

QY 453 GCGCTGTTTCCCTTCGAGCAGCAACACTGTGATATGATGTTTCGCTAGCTGACATATGA 512

Db 736 CACTTCTTCCCTTCGAGCAGCAACACTGTGATATGATGTTTGGTCTCTGAGCTTATGA 795

QY 513 CGGCAATCAGTTGGATCTGGTGTCTAAAAGATGAGGCGGCGGATCTATCGGACTTCAT 572

Db 796 CAAAGCCCAAGATCGACCTGGAGCAGATGAGCAGACTGTGG--ACCTGAAGGACTACTG 852

QY 573 AACAAATGGGAGTGTGATCTTAATAGGAATGCCAGGCAAAAGAAACACAAATTAACATCGC 632

Db 853 GGAGAGCGCGAGTGGGCCATCTGCAATGCCACGGGCACTTACAAACAGCAAGAGTACGA 912

QY 633 GTGCTGCCCGGAGCCTAGTGGAGTCACTTCCATCCATCATGATAAGAAAGACGACCTT 692

Db 913 CTGCTGGCGGAGATCTACCCGAGCTCACCTACGCTTCTGCTATCCGCGGCTGCGGCT 972

QY 693 GTACTACTTCTTCAACCTGATCGTCCGCTGCTGATCTCATGATGGCACTCTCTCGG 752

Db 973 CTTCTACACCTCAACCTCATCATCCCTGCTGCTCATCTCTCTCTCTCTCTCTCTCTCT 1032

QY 753 CTTTACACTGCCACAGACTCCGGAGAGAACTCACACTTGGAGTCACTATTTCTTCTATC 812

Db 1033 CTTTACCTGCGCTCCGACTCGCGGAGAAATCACTGCTGCTCATTTGCTGCTGCTGTC 1092

QY 813 GCTGACGCTGTTCTCTCAACCTGGTGGTGGAGACCTGCCACAGGCTCTCGAGCTATCC 872

Db 1093 ACTCACCGTCTTCTCTGCTGCTCATCTGAGATCATCCGCTCCACTCTGCTGGTCACTCC 1152

QY 873 CTTGTAGGAGAGTACTTCAATTTGATCATGTTTCAATGTTAGTGGTCTCTGTTGGTACTGAC 932

Db 1153 GCTCATCGCGAGTACCTGCTGTTTCAACATGATCTTCTGCTACCTGCTCATCATCAC 1212

QY 933 TGTGTGTGTTCAATTTACCACTCGAACAGCTGATATGATGAAATGCCACAGTGGAT 992

Db 1213 CGTCTTGTGTTCAATTTGCAACCGCTCCCGGAGCAACCAACCATGCCCACTGGGT 1272

QY 993 AAATCAGTATTCTTCAATGTTTCCCTGATGATCTCGAATGTCGAGGCC 1043

Db 1273 GCGGGGGGCCCTTCTGGGCTGTGTCGCCGCTTCTGATGAACCGGCC 1323

```
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 13, 2001 this sequence version replaced gi:12793792.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7646.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DD007CH03QP1&cluster=7646.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0DD007CH03QP1.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD007YP05"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 20.0%; Score 301.2; DB 9; Length 1201;
Best Local Similarity 60.2%; Pred. No. 1.6e-64;
Matches 537; Conservative 10; Mismatches 333; Indels 12; Gaps 3;
QY 65 ACAGAGAGAGACTCTCGAAGCGCTGTCGCGAACTACACACCCCTGGAGCGACCGGTGG 124
Db 197 WCCAGAGGAGCTTTTACAGGAGCTGTTCAGAACTACAACTCCCTTGGAGAGCCCGTGG 256
QY 125 CCAACGAGACGCAACCGCTAGAGGTTCAGGTTTCGGCTTGACCTTGCAGCAATCATTTGACG 184
Db 257 CCAATGACTCGCAACCACTACCGTCTACTTCTCCCTGAGCTCTCGCATCATGACG 316
QY 185 TGGACGAGAGATCAACTACTTATTAACCAATATATGGCTGTGCTGGAGTGAATGACT 244
Db 317 TGGATGAGAAGAACCAAGTTTAAACCACCAACATTTGGCTGCAAAATGCTTGGACAGATC 376
QY 245 ACAACCTGAGGTGGAACGACGAGTATGGCGGGGTCAAGGACCTCAGGATCACGCCCA 304
Db 377 ACTATTACAGTGAATGTGCAGAAATATCCAGGGGTGAAGACTGTTCGTTTCCCAAGATG 436
QY 305 ACAAGTTGTGGAAGCGGACGTCCTTATGTATATAGTGTCTGACGAGGGTTTTCAGCGGA 364
Db 437 GCCAGATTGTGGAACACGACATTTCTTCTATTAACAGTGTCTGATGAGCGCTTTGACGCCA 496
QY 365 CCTACAGACCAACGTTGGTGTGTCAGAAAGCGGCGGAGTTCCTGTATCGTGCACCTGGCA 424
Db 497 CATTCACACTAAACGTTGTTGGTGAATTTCTTCTGGGCAATTCGCCAGTTCCTCCAGGCA 556
QY 425 TATTCAAGAGCATCAAGATGGACATCGCTGGTTTTCCTTCGACGACCAACACTGTG 484
Db 557 TATTCAAGAGTTCCTGCTACATCGATGACGCTGTGTTTCCCTTGTATGTGACACTGCA 616
QY 485 ATATGAAGTTGGTAGTGGACATATGACGGCAATCAGTTGGATCTGGTCTATAAAGATG 544
Db 617 AACTGAAGTTGGGTCTCTGCTTACGGAGGCTGGTCCCTTGGATCTGCAGATCCAGGA-- 674
QY 545 AGCAGGCGCGATCTATCGACTTCTATPACAAATGGGAGTGATGTATCTAATAGGAATGC 604
Db 675 -----GGCAGATATCAGTGGCTATATCCCCAATGGAATGGGACCTAGTGGGAATCC 727
```

```
QY 605 CAGGCAAAAGAACACAATAACATACGCTGCTGCCCCCGAGCCCTAGCTGGAGCTCACCT 664
Db 728 CCGGCAAGAGGAGTGAAGGTTCTATGAGTCTKTCGAAAGAGAGCCCTACCCGATGTACCT 787
QY 665 TCACCATCATGATGAAGAAGACGAACCTTGTACTACTTCTTCAACCTGATCGTCCCGTGG 724
Db 788 TCACAGTGCACATGCCCGCAGGACGCTCTACTATGCGCTCAACCTGCTGATCCCTGTG 847
QY 725 TGCTGATCATGATGAGGACCTCTCGGCTTACACTGCGCACACGAGACTCCGAG--AGAA 782
Db 848 TGCTCATCTCCGCCCTCGCCCTGCTGTTCTCTGCTGAGATTCGCGGGGAGAG 907
QY 783 ACTCACACTTGGAGTCACTATTCTTCTATCGCTGACGGTGTCTCAACTGCTAGCCGA 842
Db 908 ATTTCCTGGGATTAACAGTATTCTCTCTTACCGTCTTCTGCTGCTGCTGGCTGA 967
QY 843 GACCTCGCCACAGGCTCTCGAGCGCTATCCCGCTGTTAGGAGCGTACTTCAATTGCATCAT 902
Db 968 GATCATGCCCGCAACATCCGATTTCGGTACCAATTGAWAGCCCACTTCTTCCGACGACCAT 1027
QY 903 GTTCATGGTAGCGTGTCTGTGGTACTGACTGTGGTGGTACTCAATTACCAC 954
Db 1028 GATCATCGT-GGGCTTTGGGTGGTTGKRMGGTGTGCTGCTGACGTACAC 1078

RESULT 5
AY402873 1436 bp DNA linear GSS 15-DEC-2003
LOCUS Homo sapiens CHRNA3 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY402873
VERSION AY402873.1 GI:39758856
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1436)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1436)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment
FEATURES
source
1..1436
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1_>1436
/gene="CHRNA3"
/locus_tag="HCML1369"

ORIGIN
Query Match 19.6%; Score 294.8; DB 29; Length 1436;
Best Local Similarity 57.1%; Pred. No. 7e-63;
Matches 558; Conservative 0; Mismatches 417; Indels 3; Gaps 1;
QY 67 GAGAAGAGACTCTGAAACGCTGCTGGCGAACTACAAACCCCTGGAGGACCGGTGGCC 126
```


Db 21 GAGCACCCTCTATTGTGAGCGGCTGTTTGAAGATTACAAATGAGATCATCGGCGCTGTGCC 80
QY 127 AACGAGAGCGAACCCCTAGAGGTGAGGTTCGGCTTGACCTTCGACAAATCATTTGACGTG 186
Db 81 AACGTGTCTGACCCAGTCATCATCCATTTCGAGGTGTCATGCTCTCAGCTGGTGAAGGTG 140
QY 187 GACGAGAGAACTACTATTATACCAATATATATGCTGTCTGTTGAGTGGNATGACTAC 246
Db 141 GATGAAGTAACACGATCATGGAGACCAACCTGTGGCTCAAGCAATCTGGAATGACTAC 200
QY 247 AACCTGAGGTGGAACGACGAGCGAGTATGCGGGGTCAAGGACCTCAGGATCACGCCCAAC 306
Db 201 AAGCTGAAGTGAACCCCTCTGACTATGTTGGGCGAGATTATGCTGCTCCTCCACAG 260
QY 307 AAGTTGTGAAGCCGAGCTCTTATGTATATATATGTTGCTGACGAGGTTTTCACGGGACC 366
Db 261 AAGATCTGGAAGCCAGACATTTGCTGTATATCAATGCTGTTGGGGAATTCAGGTGGAC 320
QY 367 TACCAGACCAACGTGTGTCAGAGCGCGCGCAGTTGCTGTACGTGCGCACTCGGCATA 426
Db 321 GACAAGACCAAGCCTTACTCAAGTACACTGGGAGGTGACTTGGATACCTCGGCCCATC 380
QY 427 TTCAGAGCACATGCAAGATGAGATCGCGTGGTTTCCCTTCGACGACCAACACTGTGAT 486
Db 381 TTTAAGAGCTCCTGTAAATCGAGCTGACCTACTTCCCGTTTGATTACCAAACTGTACC 440
QY 487 ATGAAGTTCGTGAGTGCATATGACGGAATCAGTTGGATCTGGTGTAAAGATGAG 546
Db 441 ATGAAGTTCGTGTCGTGCTTACGATAAGGCGAAATCGATCTGCTGCTGA---TCGCG 497
QY 547 GCAGCGCGGATCTATCGGACTTCAACAAATGGGAGTGTATCTAATGAATGCCA 606
Db 498 TCTTCATGAACCTCAAGACTATTGGGAGCGCGGAGTGGGCCATCATCAAGGCCCA 557
QY 607 GCGAAAAAGAACACATAACATACGCGTGTCTGCCCGAGCCCTACGTGACCTCACCTTC 666
Db 558 GGCTACAAACGACATCAAGTACAACTGCTCGGAGGAGATCTACCCGACATCACATAC 617
QY 667 ACCATCATGATGAAGAGAGAACCTTGTACTACTTCTCAACCTGANTGCTCCGTGGGTG 726
Db 618 TCGCTGTACATCGGCGGCTGCGCTTGTCTACACCATCAACCTCATCTCCCTGCGCTG 677
QY 727 CTGATCTCATGATGCGACTCTCGGCTTCACTGCGCACAGACTCCGGAGAGAACTC 786
Db 738 ACCCTGTGCATTTCTGCTCCTCTCCTCGACGCGTGTCTCTCTGTGATCACTGAGACC 797
QY 847 CTGCCACAGGTCTCGACGCTATCCCGCTGTAGGAGCGTACTTCAATTCGATCATGTTC 906
Db 798 ATCCCTTCCACCTCGCTGCTATCCCGCTGTAGGAGTACTCTCTGTTCACCATGATT 857
QY 907 ATGGTAGCGTCTCTGTGTGCTGACTGTGGTGGTACTCAATTAACACCATCGAACAGCT 966
Db 858 TTGTGAACCTGTGCTCATGCTATCAACGCTCTCGTGTCAAGTGCATGACAGAACCCCG 917
QY 967 GATATACATGAATGCCAGTGGATAAATCAGTATTCTCTCAATGGTTGCCATGGATA 1026
Db 918 ACGACACACACAATGCCCTCATGGTGAAGACTGTATTCTTGAACCTGCTCCCGAGGTC 977
QY 1027 CTGCGAATGTCAGGGCA 1044
Db 978 ATGTTTCATGACGACGGCA 995

RESULT 6

AK034228

LOCUS

DEFINITION

2940 bp mRNA linear HTC 18-SEP-2003
Mus musculus adult male diencephalon cDNA, RIKEN full-length
enriched library, clone:933016516 product:cholinergic receptor,
nicotinic, alpha polypeptide 4, full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK034228
AK034228.1 GI:26329798
HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

REFERENCE

AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

REFERENCE

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumi, T., Tashiro, H., Itoh, M.,
Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, S., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multipillarary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

REFERENCE

AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the

FANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

REFERENCE

AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 2940)

REFERENCE

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,

Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,

Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,

Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,

Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,

Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,

Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaki-Akashira, S.,

Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,

Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,

Fax: 81-45-503-9216]

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/
 Location/Qualifiers
 1. 2940
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:9330165116"
 /db_xref="MGI:2398424"
 /db_xref="taxon:10090"
 /clone="9330165116"
 /sex="male"
 /tissue_type="diencephalon"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 116. 2005
 /note="unnamed protein product; cholinergic receptor, nicotinic, alpha polypeptide 4 (MGD|MGI:87888, GB|NM_015730, evidence: BLASTN, 99%, match=1946) putative"
 /codon_start=1
 /protein_id="BAC28638.1"
 /db_xref="GI:26329799"
 /translation="MEIGSGAPPLPILPLLLLTGLLPASSHLETHAEERLLK RFGNGKNSRFPANISDVLRFGLSIAQLDVIDEKNQMTINVMKQNDYKLRW DPGDYENVTSIRIPSELIWRPDIYLYNADGDFAVTHLKAHLFYDGRVQTPPALYK SSCIDVTFFPDQDQTKMFGSYDKAKIDLVSMSRVDQLDFWESGEWIVDAVG TYNTRYECABIPDIYVAFIIRRLPLFTINLIIPCLLISCLTVLFLPSECGEK VTLCSVLSLAVFLLIETIIPSTSLVPLIGELLYFTWIFVLSIVITVFLNVHH RSPRHTWPAWVRVLDIVPRLLFMKRSVVDNCRLLIESHMRMANAPRPWPPEP ERLGIDICNQLSPAPFCNMDTAVETQPCRSFSPKVPDLKTSVEKSPCPSPG SCPPNSGAPVLLIKARSLSQVPSQBAEGSIRCSRSIQYCVSDGAASLPTSK PTGSPASLRTFSQVPSQBAEGSIRCSRSIQYCVSDGAASLPTSK TRAVEGVQYIADHLKAEIDFVSKEDWYVAVVIDRIFLWMPFIIVCLLTGVLFLPPW LAGMI"

ORIGIN

Query Match 19.5%; Score 293.2; DB 11; Length 2940;
 Best Local Similarity 55.9%; Pred. No. 2.5e-62;
 Matches 578; Conservative 0; Mismatches 453; Indels 3; Gaps 1;
 QY 11 TGTGGCGGCTTGGCGCTTGTGCTGCGGATCGGAGCAAGGTCTCTCAGAGA 70
 Db 174 TCTTAGGACCGGCTTCTGCTGCTAGGACGACATAGAGACCGGGCCATCGCGAGG 233
 QY 71 AGAGACTCTGAAACGGTGTGCTGGCGAACTTAAACACCTCGAGCGACCGGTGGCCAAAG 130
 Db 234 AGCGGCTCTGAAGAGACTCTTCTGGCTACACAAAGTGTCTCGCCAGTAGCCAAATA 293
 QY 131 AGAGCAACCGCTAGAGGTGAGTTCGGCTTGACCTTGACGCAAAATCATGTAGCTGGAG 190
 Db 294 TCTCAGATGTGTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 353
 QY 191 AGAAGAATCAACTACTTATAACCAATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 250
 Db 354 AGAANAACAGATGATGACGACCAACGCTGTGGTGAAGCAGAGTGGCATGACTACAAAC 413
 QY 251 TGAGGTGGAACGACGAGTATGCGGGGTCAAGGACCTCAAGGATCAGGACCGCCCAACAAGT 310
 Db 414 TCGCGTGGGACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 473
 QY 311 TGTGNAAGCGGACGCTCTTATGATATAGTGTGCTGACGAGGGTTTGTAGGGACCTAC 370
 Db 474 TCTGGAGGCTGACATCGCTCTTCTACAAACGCGGACGGGACCTTTGCAAGTCAACCCACC 533
 QY 371 AGACCAACGCTGTGTGTCAGAGCGGCGGAGTGTGCTGTACGTGCCACCTGGCATATCA 430
 Db 534 TAACCAAGCCACCTGTCTATGATGGCGGTGTGCGATGGACACCCCGGCATCTATA 593
 QY 431 AGAGCATGCAAGATGGAATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 490
 Db 594 AGAGCTCTGCAATCGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 653
 QY 491 AGTTCGGTAGCTGGACATATGACGCAATCAGTTGGATCTGTGCTGCTGCTGCTGCTGCTG 550

Db 654 AGTTTGGGTCTGGAGCTTACGACCAAGGCTTACCTTGGTGGATGACATGACAGCGGTG 713
 QY 551 GGGCGCATCTATCGGACTTCTATCAAAATGGGAGTGGTATCTAAATAGGATGCCAGGCA 610
 Db 714 TGGACCAACT--GGACTTCTGGGAAAGTGGGAGTGGGTCATTTGGATGCGCGTGGGCA 770
 QY 611 AAAAGAACCAATAACATACGCTGTGCGCCGAGCCCTACGTGGAGCTGACCTTCCACCA 670
 Db 771 CCTACCAACCCAGGAGTATGATGTGTGCGGAGATCTATCTGATCATCCTACCTACGCT 830
 QY 671 TCATGATAGAGAGCAAACTTGTACTACTTCTTCAACCTGATGCTGCTGCTGCTGCTGCTG 730
 Db 831 TCATCATCCGCGACTGCGCACTGTTCTACACCATCAACCTTATCATCCCGTGGCTGCTCA 890
 QY 731 TCTCATGATGCGACTCTCTGCGCTTCCACCTGCGGAGCTGCGGAGGAACTACAC 790
 Db 891 TCTCTGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 950
 QY 791 TTGGAGTCACTATTTCTTATCGCTGACGCTGCTTCTCAACCTGCTGCTGCTGCTGCTGCTG 850
 Db 951 TGTGATCTCGGTGCTGCTTCTCTCACCGTCTTCTGCTGCTCATCACCGAGATCATCC 1010
 QY 851 CACAGGTCTCCGAGCGTATCCCGCTGTTAGGAGCGTCTCAATTCGATCATGCTTCATGG 910
 Db 1011 CGTCCACCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1070
 QY 911 TAGCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 970
 Db 1071 TCACCTCTCCATTTGTCATCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1130
 QY 971 TCATGAAATGCGACAGTGGATAAAATCAGTATTCCTACATGTTGGCTGCTGCTGCTGCTGCT 1030
 Db 1131 CACACACCATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1190
 QY 1031 GAATGTCGAGGCA 1044
 Db 1191 TCATGAAGCGCCA 1204
 RESULT 7
 AK083157
 LOCUS
 DEFINITION
 Mus musculus adult male hippocampus cDNA, RIKEN full-length enriched library, clone:C630019M18 product:cholinergic receptor, nicotinic, alpha polypeptide 4, full insert sequence.
 AK083157
 VERSION
 AK083157.1 GI:26350296
 HTG; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1
 Carninci, P. and Hayashizaki, Y.
 TITLE
 High-efficiency full-length cDNA cloning
 JOURNAL
 Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE
 99279253
 PUBMED
 10349636
 REFERENCE
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL
 Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE
 20499374
 PUBMED
 11042159
 REFERENCE
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanishi, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
 Riken integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 PUBMED
 11076861

REFERENCES

4
 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 3230)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
 Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/

FEATURES

Location/Qualifiers
 1..3230
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CS7BL/6J"
 /db_xref="FANTOM_DB:C630019M18"
 /db_xref="MGI:2417536"
 /db_xref="taxon:10090"
 /clone="C630019M18"
 /sex="male"
 /tissue_type="hippocampus"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 36..1925
 /note="unnamed protein product; cholinergic receptor, nicotinic, alpha polypeptide 4 (MGI:MGI:87888, GB|NM_015730, evidence: BLASTN, 99%, match=1946) putative"
 /codon_start=1
 /protein_id="BAC38788.1"
 /db_xref="GI:26350297"

CDS

/translation="MEIGSGAPPPLLLPILLIGTLLPASSHETRAHEERLLK RLPSGNKSRVNPISLWVIRFVGLNADGDFAVTLTKAHLFDYGVQWTPPAIYK SSCSDVTFPPDQNCMTKFGSYTDKADLDVSMHSRVDQLDFWESGEWIVDAVG TYNTRKYCCAEIYDITVAFIIRLPLFYTLINLIIPCLISLTLYLPSGEGX

ORIGIN

Query Match 19.5%; Score 293.2; DB 11; Length 3230;
 Best Local Similarity 55.9%; Pred. No. 2.7e-62;
 Matches 578; Conservative 0; Mismatches 453; Indels 3; Gaps 1;
 LAGMI"
 VTLICISVLISLTVFLLILTEIIPSTSLVILPILGEYLLFTWIFVLSIVITVFNVHH
 RSPRTMPAVRVRVFLDIIVRLFLMKRPSVWKNCRLLIESKMANARFPWPEPS
 EPGILGDIQNGLSPPATFTCRMDTAVETQPTCRSPSHKVPDLKTEVEKASPCPSFG
 SCHPPNSSGAPVLIKARSLSVQHVPSSEAEGSIRCSRPSIOYCVSDGAASITSEK
 PTGSPASLAKTRPSQLPVSQTSCKCTCKEPPSPVITVLKAGTKAPPHLPLSPAL
 TRAVEGVQVIADHLKAEIDTDFSVKEDWKVYAMVIDRIFLWMFIIIVCLLGTGVLFPWP
 LAGMI"
 QY 11 TGTGGCGCGCTTGGCGCTGCTGGCTTTGCTGCGCTTATCGGAGCAAGGTCTTCACGAGA 70
 Db 94 TCTTAGGGACCGGCTCTTGGCTGTAGCAGCCACATAGAGACCGCGGCCCATGCGGAGG 153
 QY 71 AGAGACTCCTGAACGGCTTGGCTGGCGAACTACACACCTGGAGCGACCGGTGGCCACG 130
 Db 154 AGCGGCTCCTGAAGAGACTTCTCTCTGGCTTACACCAAGTGGTCTCGGCCAGTAGCCAATA 213
 QY 131 AGAGCGAACCGCTAGAGGTGAGGTTCGGCTTGACCTTGCAGCAAAATCATTTGACGTGACG 190
 Db 214 TCTCAGATGGTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 273
 QY 191 AGAGAACTCAACTACTATTATAACCAATATATGCTGCTGCTGGAGTGGAAATGACTACAAC 250
 Db 274 AGAAAAACAGATGATGACGACCAACGTTGGGTGAAGCAGGAGTGGCATGACTACAAAC 333
 QY 251 TGAGGTGGAAACACAGCAGGAGTATGGCGGCTCAAGCACTCAGGATCAGCCCCAACAGT 310
 Db 334 TGGCTGGACCGCTGGTGACTAGAGATGTCACTCCATCCGATCCCATCTGGAATCA 393
 QY 311 TGTGAAGCGCGACGCTCTTATGATAATAGTGTGACGAGGCTTTTACGCGGACCTTACC 370
 Db 394 TCTGGAGGCTGACATCGTCTTACACACACGCGACGCGGAGCTTTGAGTCACTCACCACC 453
 QY 371 AGACCAACGTGTGTGTGAGAGCGCGGAGTTGCTGTACCTGGCCACTTGGCATATTCA 430
 Db 454 TAAACAAAGCCACCTGTTCTATGATGGCGGTGTGCAGTGGACACCCCGGCGCATATA 513
 QY 431 AGAGCACATGCAAGATGGACATCGGTGGTTTCCCTTGCACACCAACACTGTGATAGA 490
 Db 514 AGAGTCTCTGACAGATCGACGTCACTTCTTCCCTTGCACGAGAACTGTACCATGA 573
 QY 491 AGTTCGTAGCTGGACATATGATGACGCAATCAGTGTGATCTGCTGTCTTAAAGATAGGCGAG 550
 Db 574 AGTTTGGGTCTGGACCTACGACAGGCCAAAGTGAATGACTTGGTGGATGACACGCCGTG 633
 QY 551 GCGGCGATCTATCGGACTTCTATAACAAATGGGAGTGGTATCTAATAGGAATGCCAGGCA 610
 Db 634 TGGACCAACT---GGACTTCTGGGAAAGTGGGAGTGGGTCAATTGTGGATGCGGTGGCA 690
 QY 611 AAAAGAACACATAACATACGCTGCTGCCCGAGCCCTTACGTGAGCTCACTTCACTCA 670
 Db 691 CCTACACACAGGAAGTATGAATCTGTGCCGAGATCTATCTCTGACATCACTACGCT 750
 QY 671 TCATGATAAGAAAGACCAACCTTGTACTCTTCTCAACCTGATCTCGCTCGGTGGTCTGA 730
 Db 751 TCATCATCGCGGACTGCCACTGTTCTACACCATCACTTATCATCCGCTGCTGCTCA 810
 QY 731 TCTCATGATGGCACTCTCTCGGCTTCACTGCCACCAAGACTCCGAGAGAACTCACAC 790
 Db 811 TCTCTGCTCACCCTGCTGCTTCTATCTGCTTCTTCTCTCCTCTCTCATCCAGGATCATCC 870
 QY 791 TTGGAGTCACTATTCTTCTATGCTGACGCTGCTTCTCAACCTGATCTCGCTCGGTGGT 850
 Db 871 TGTGATCTCGGTGCTGCTTCTCTCACCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 930
 QY 851 CACAGGTCTCGGAGCTATCCCTCTTGTAGGAGACGCTTCAATTGATCATGTTTCATGG 910
 Db 931 CGTCCACCTCGTGTGATCCCGCTCATCGCGGAGTACTGCTCTTCACTGATCTCTCG 990

```
QY 911 TAGCGTCGCTCTGTGGTACTGACTGTGGTGTACTCAATTACCACCATCGAAGCTGATA 970
Db 991 TCACCTCTCCATTGTCATCAGGTCTTCTGCTCTCAATGTACACACCGCTCACCAGCA 1050
QY 971 TACATGAATGCGACAGTGGATAAATCAGTATCTTCAATGGTGGCATGATCTGC 1030
Db 1051 CACACACCATGCGCGCTGGTGGCAGAGTCTTCTCGGACATGTGGCCCGCTCTCTCT 1110
QY 1031 GRATGTCGAGGCCA 1044
Db 1111 TCATGAAGCGCCCA 1124

RESULT 8
AY402874 1436 bp DNA linear GSS 15-DEC-2003
LOCUS Pan troglodytes CHRNA3 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY402874
VERSION AY402874.1 GI:39758857
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 1436)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene tricos
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1436)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..1436
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>1436
/gene="CHRNA3"
/locus_tag="HCML1369"

ORIGIN
Query Match 19.1%; Score 286.8; DB 29; Length 1436;
Best Local Similarity 55.6%; Pred. No. 7.3e-61;
Matches 544; Conservative 0; Mismatches 431; Indels 3; Gaps 1;

QY 67 GAGAGAGACTCCTGAACGGTCTCTCGGCACTACACACCCCTGGAGCGCGTGGCC 126
Db 21 GAGCAGCGTCTATTGTAGCGGCTGTTTGAAGATTACAAATGAGATCATCCGGCGCTGTGGCC 80
QY 127 AACGAGAGCAACCGCTAGAGGTGAGGTTCGGCTTGACCTTGCAGCAAAATCATTGACGTG 186
Db 81 AACGTGTCAGCCAGTCATCATCTATTNNNGTGTCCATGTCTNNNTGGTGAAGTG 140
QY 187 GACGAGAAGATCACTACTTATAACCAATATATGGCTGTGCTGGAGTGAATGACTAC 246
Db 141 GATGAAGTAAACCGAGATCATGGAGCAACCACTNNNGCTCAAGCAAAATCTGGAATGACTAC 200
QY 247 AACCTGAGGTGGAACGACGAGGTATGTCGGGGGTCAAGGACCTCAGGATCACGCCCAAC 306
Db 201 AAGCTGAAGTGAACCCCTCTGACTATGGTGGGCGAGGTTCATGCGTCCCTGNNNN 260
```

```
QY 307 AAGTTTGGAAAGCCGACGCTCCTTATGTATAATAGTGTGCTGACGAGGTTTTCAGCGAAC 366
Db 261 NNNNNCTGGAAGCCNNACATTTGCTGTATATAAATGCTGTTGGGATTTCCAGGTGGAC 320
QY 367 TACAGACCAACGTTGGTGTGAGAGCGGCGGAGTTCCTGTACGTGCGACCTGCGATA 426
Db 321 GACAAGACCAAAAGCCTTACTCAAGTACACTGGGGAGGTGACTTGGATACCTCCGGGCATC 380
QY 427 TTTCAAGACACATGCAAGATGACATCGCGTGTGTTTCCCTTCGACACCAACAACCTGTGAT 486
Db 381 TTTAAGAGCTCTCTGCAAAATCGACGTGACTCTTCCCGTTTGGATTACCAAACTGTACC 440
QY 487 ATGAAGTTTCGTAGCTGGACATATGACGGCAATCAGTTGGATCTGTGTCTTAAAGATGAG 546
Db 441 ATGAAGTTTCGTCTGCTGCTTACGATAAGGGCGAAATCGATCTGCTCTGA---TCGGC 497
QY 547 GAGGGCGGCACTATCGGACTTCATAACAATGGGAGTGGTATCTAATAGAAATGCCA 606
Db 498 TCTTCCATGAACCTCAAGGACTATTGGGAGAGCGGAGTGGGCCATCATCAAGGCCCA 557
QY 607 GGCAAAAGAACACAATAACATACGCGTGTGCTGCCCGAGCCCTACGTGGACGTCAACCTTC 666
Db 558 GGTCTACAAACAGACATCAAGTACAACTGCTNNNAGAGATCTACCCCGACATCACATAC 617
QY 667 ACCATCATGATGAAGAAGACGAACTTGTACTACTTCTTCAACCTGATGCTCCCGTGGTG 726
Db 618 TCGCTGTACATCCGGCGCTCGCCTTGTCTACACCATCAACCTCATCTCCCTGCTGCTG 677
QY 727 CTGATCTCATGATGACATCTCCCTGGCTTCACACTGCCACGACTCCGGAGAGAATC 786
Db 678 CTCATCTCTCTCTCTACTGTGCTGCTTCTACTGCTCCCTCCGACTCGGTGAGAAGTG 737
QY 787 ACACTTGGAGTCACTATTCTTCTATCGTACCGGTGTTCTCAACCTGGTAGCGGAGACC 846
Db 738 ACCCTGTGATTTCTGCTCTCTCTCTCTCTGACGGTGTCTCTCTGGTGTGATCACTGAGACC 797
QY 847 CTGCCACAGGTCTCCGACGCTATCCCGCTGTTTAGGACAGTACTTCAATTCGATCATGTTTC 906
Db 798 ATCCCTTCCACCTCGCTGCTGCTATCCCGCTGATTGGAGAGTACCTCTCTGTTTCCACATGATT 857
QY 907 ATGGTAGCTGCTGCTGTGGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
Db 858 TTTGTACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 917
QY 967 GATATACATGAATGCCACAGTGGATAAAATCAGTATTCCTACAATGTTGCTCATGATA 1026
Db 918 ACACACACACATGACCTCATGGTCAAGACTGTATTCTTGAACCTGCTCCCGAGGTC 977
QY 1027 CTCGGAATGTCGAGGCCA 1044
Db 978 ATGTTTCATGACGAGGCCA 995

RESULT 9
AK081254 3483 bp mRNA linear HTC 20-SEP-2003
LOCUS Mus musculus adult male corpus striatum cDNA, RIKEN full-length
DEFINITION enriched library, clone:030030P04 product:cholinergic receptor,
nicotinic, beta polypeptide 2 (neuronal), full insert sequence.
ACCESSION AK081254
VERSION AK081254.1 GI:26099790
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
```

```

REFERENCE
AUTHORS
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
MEDLINE
Genome Res. 10 (10), 1617-1630 (2000)
PUBMED
20499374
REFERENCE
AUTHORS
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL
MEDLINE
Genome Res. 10 (11), 1757-1771 (2000)
PUBMED
20530913
REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
MEDLINE
Nature 409, 695-690 (2001)
PUBMED
11076861
REFERENCE
AUTHORS
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL
MEDLINE
Nature 420, 563-573 (2002)
PUBMED
12045159
REFERENCE
AUTHORS
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tanaka,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
TITLE
Direct Submission
JOURNAL
MEDLINE
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES
source
Location/Qualifiers
1..3483
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:C030030P04"
/db_xref="MGI:2413538"
/db_xref="taxon:10090"
/clone="C030030P04"
/sex="male"
/tissue_type="corpus striatum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"

```

```

/dev_stage="adult"
1..3483
/notes="cholinergic receptor, nicotinic, beta polypeptide 2
(neuronal) (MGID|GI:87891, GI|NM_009602, evidence: BLASTN,
99%, match=1498)"
ORIGIN
Query Match 18.7%; Score 281.6; DB 11; Length 3483;
Best Local Similarity 57.7%; Pred. No. 2.3e-59;
Matches 547; Conservative 0; Mismatches 389; Indels 12; Gaps 2;
QY 99 CTACAAACCCCTGGAGCGAGCGGTGGCCCAACAGAGAGCGTAGAGGTGAGGTTCGG 158
DB 326 CTATATAAAGCTGATCCGTCAGCTACTAATGGCTCTGAGCTGGTACGTGTACAGCTCAT 385
QY 159 CTTGACCTTTGACGAAATCATTCAGCTGACGAGAGAGAAATCACTACTTATTATACCAATAT 218
DB 386 GGTATCATTTGGCACAGCTCATCAGTGTGCACGAGCGGAGCAGATCATGACCAACGCT 445
QY 219 ATGGCTGTCTGGAGTGAATGACTACAACTGAGGTGGAACGAGCAGCGAGTATGGCGG 278
DB 446 CTGGCTGACCCAGGAGTGGGAAGATTATCGCTTCACATGGAAGCCTGAGGATTCGACAA 505
QY 279 GGTCAAGGACCTCAGGATCAGCCCAACAAAGTTGTGGAAGCGGAGCTCTTATGTATAA 338
DB 506 TATGAAGAAGTCCGACTCCCTTCTTAAGCAGATCTGGCTCCAGATGTGGTCTTATACAA 565
QY 339 TAGTGTCTGACGAGGGTTTTGACGGGACCTACAGACCAACGTTGGTGGTTCAGAAAGCGCGG 398
DB 566 CAATGCTGACGGCATGTACGAAGTCTCTTCTTATTCATGCTGTGTCTCTCTATGATGG 625
QY 399 CAGTTCCTGTAGTGGCCACCTGGGCATATTTAAGAGCAGCATGCAAGATGAGCATTCGGGTG 458
DB 626 CAGCATCTTTTGGCTACCGCTGCCATCTACAGAGCGCATGCAAGATTTGAGGTGAAGCA 685
QY 459 GTTTCCTTCGAGGACCAACACTGTGATGATGAAGTTGGTGGTGGTACGACATATCAGCGCAA 518
DB 686 CTTCCCATTTGACACGAGACTGCAACATGAAGTTCCGCTCTGGACCTTACGACCGCAC 745
QY 519 TCAGTTGGATCTGGTCTAAAGATGAGGAGCGGCGGCGATCTATCGGACTTCTATAACAA 578
DB 746 TGAGATTGACCTGGTCTCAAAAGCGATGTGGCGAC--CTGGAGCGACTTTCACACCTAG 802
QY 579 TGGGAGTGGTATCTAATAGGATGCCAGGCAAAAGAACACATACATACCGCTGCTG 638
DB 803 TGGGAGTGGGACATCATCGCATGCGAGCGGCGAGCAACGAGAACCCAGACGACTCC-- 860
QY 639 CCGGAGCGCTACGTGGAGCTCACCCTTACCACATCATGATAAGAGAGCAACCTTGTACTA 698
DB 861 -----ACTAGTGGACATCACCTACGACTTCATTCGTCGCAACCGCTCTCTTA 913
QY 699 CTTCTTCAACCTGATCGTCCGCTGGTGTGATCTCATCGATGGCACTCTCTGGGCTTAC 758
DB 914 CACCATCAACCTCATCATCCCTGGCTACTCATCCTCGTGGCCATCCTGGTCTTCTTA 973
QY 759 ACTGCCACAGACTCCGGAGAGAACTCAGACTTGGAGTCACTATTTCTTATCGCTGAC 818
DB 974 CTTGCCCTCAGACTGTGGTGAAGATGACACTTTGTATTTCTGTCTGTGGCGCTAC 1033
QY 819 GGTGTTCTTCAACCTGGTAGCGGAGACCTGCCACAGGCTCTCGACGCTATCCCTCTGT 878
DB 1034 GGTGTTCTTCTGTGCTCATCTCAAGATTGTGCTCCACCTCTCCAGAGTACCGTGT 1093
QY 879 AGGAGCGTACTTCAATTGTCATCATGTTATGGTGTAGCGTCTGTGTGGTACTGACTGTGT 938
DB 1094 GGGAAAGTACCTCATGTTTCAACATGGTGTAGTACCTTTCTCATGCTCCTAGCGGTG 1153
QY 939 GGTACTCAATTACCACCATCGAACAGCTGATATATGAAATGCCACAGTGGATGATAAATC 998
DB 1154 TGTGCTCAATGTGCACACCGTTCGCTTACCGACACCATGGCGCCCTGGGTCAAGGT 1213
QY 999 AGTATTCTTACATGTTGCCATGGATCTGCGAATGTGAGGCCAGG 1046

```


QY 519 TCAGTTGGATCGGTCTAAGAGATGAGCGCGCGATCTATCGGACTTCATACAAA 578
 Db 742 TGAGATTGACCTGGTCTCAAGAGCATGTGGCCAGC---CTGGAGCATTCACACTAG 798
 QY 579 TGGGGAGTGGTATCTAATAGGAATCCAGGCAAAAAGAACACAAATACATACGCGTGGT 638
 Db 799 TGGGGAGTGGGACATCATCGCATCGCCAGCGGACGACGAGACCCAGACGACTCC-- 856
 QY 639 CCCGAGCCCTAGTGGAGTGCACCTTACCATCATGATGAAGAGACGACCTTGTACTA 698
 Db 857 -----ACCTAGCTGGACATCACCTACGACTTCATCATTCGTCGCAACCGCTCTTCTA 909
 QY 699 CTTCTTCAACCTGATCGTCCCGTGGTGTCTCATCGATGCGGCACTCTCGGCTTAC 758
 Db 910 CACCATCACTCATCATCCCTGGCTACTCATCATCGCTGGCCATCCTGCTTCTTA 969
 QY 759 ACTGCCACAGACTCCGGAGAGAAACTCACACTTGGAGTCACTATTCTTCTATCGCTGAC 818
 Db 970 CTGCGCCTCAGACTGTGGTGAAGATGACACTTTGTTTCTGTCTGCTGGCGCTAC 1029
 QY 819 GGTGTTCTCAACCTGGTGGAGACCTCGCACAGGTCCTGGAGCGCTATCCCGCTGTT 878
 Db 1030 GGTGTTCTCGTCTCATCTCCAAAGATTGCGCTCCACCTCCCTCGAAGTACCGTGT 1089
 QY 879 AGGAGCTACTTCAATTGTCATCATGTTCTATGGTGGTGTCTGTGCTACTGACTGTGT 938
 Db 1090 GGAAGTACCTCATGTTTACCATGGTGTAGTACCTTCTCATGCTCAGTGGTG 1149
 QY 939 GGTACTCAATTACCATCGAACAGCTGATATACATGAATGCCACAGTGGATAAATC 998
 Db 1150 TGTGCTCAATGTGCACACCGTTCGCTACCAGCACCATGGCGCTGGTCAAGT 1209
 QY 999 AGTATTCCTCAATGTTGCCATGGATACCTGGAATGTGAGGCCAGG 1046
 Db 1210 GGTCTTCTGGAGAGCTGCCACCTCCTCTTCTTCTGACGACGCCAGC 1257

RESULT 11
 AK051742
 LOCUS
 DEFINITION
 Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length
 enriched library, clone: D130070121 product: cholinergic receptor,
 nicotinic, beta polypeptide 2 (neuronal), full insert sequence.
 AK051742
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus
 HTC; CAP trapper.
 Mus musculus (house mouse)
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 2
 20499374
 MEDLINE
 PUBMED
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yonekura, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format

JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

TITLE JOURNAL REFERENCE AUTHORS

TITLE JOURNAL REFERENCE AUTHORS

TITLE JOURNAL

COMMENT

FEATURES source

1. 4046
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:D130070121"
 /db_xref="MGI:2420100"
 /db_xref="taxon:10090"
 /clone="D130070121"
 /tissue_type="spinal ganglion"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="12 days embryo"
 209. 1714

CDS

/note="unnamed protein product; cholinergic receptor,
 nicotinic, beta polypeptide 2 (neuronal) (MGI:MGI:87891,
 GB|NM_009602, evidence: BLASTN, 99%, match=1498)
 putative"
 /codon_start=1
 /protein_id="BAC34749.1"
 /db_xref="GI:26342174"
 /translation="MARCSMALLFSGLLWLCGVLGTDEERLVEHLDPSSYNNK
 LIRPTNGSELVTQVWYSLAQISVHEREQIMTVMLTQEWEDYRLTWKDEDDNM
 KKVRLPSKHLPDVLVLYNADGMVEFVSNVYDGSFWLPPAIYKSKACKIEVK
 HFFPDQQTMKFRSWTYDRTEDILVLCSDVASDDFTPSGEWDIIALPGRNENPDD
 STYVDYDFIRKPLFTNLIPLVITSLAIFVLPDSCGCKTLCISVLLAL
 TVFLILSKIVPPTSLDPLVLMKFTMLVITFVSCVNLVHRSPTHTWAPW
 FVWLEKLPTLLFQQPHRCARORLRRORREGAGTLFFEGPAADPTCFVN
 PASMOGLAGAQAEPAAAGLSMGPCSGLEAREADVGRFTADHRSDDDSVREDW
 KIVAVDLFLFWVFCVFTIGMFLQPLFQNTATTFTLHSDHSAFSSK"

sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861

4
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)

5
 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

6 (bases 1 to 4046)

Adachi, J., Aizawa, K., Kimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

Location/Qualifiers

1. 4046

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:D130070121"

/db_xref="MGI:2420100"

/db_xref="taxon:10090"

/clone="D130070121"

/tissue_type="spinal ganglion"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="12 days embryo"

209. 1714

/note="unnamed protein product; cholinergic receptor,

nicotinic, beta polypeptide 2 (neuronal) (MGI:MGI:87891,

GB|NM_009602, evidence: BLASTN, 99%, match=1498)

putative"

/codon_start=1

/protein_id="BAC34749.1"

/db_xref="GI:26342174"

/translation="MARCSMALLFSGLLWLCGVLGTDEERLVEHLDPSSYNNK
 LIRPTNGSELVTQVWYSLAQISVHEREQIMTVMLTQEWEDYRLTWKDEDDNM
 KKVRLPSKHLPDVLVLYNADGMVEFVSNVYDGSFWLPPAIYKSKACKIEVK
 HFFPDQQTMKFRSWTYDRTEDILVLCSDVASDDFTPSGEWDIIALPGRNENPDD
 STYVDYDFIRKPLFTNLIPLVITSLAIFVLPDSCGCKTLCISVLLAL
 TVFLILSKIVPPTSLDPLVLMKFTMLVITFVSCVNLVHRSPTHTWAPW
 FVWLEKLPTLLFQQPHRCARORLRRORREGAGTLFFEGPAADPTCFVN
 PASMOGLAGAQAEPAAAGLSMGPCSGLEAREADVGRFTADHRSDDDSVREDW
 KIVAVDLFLFWVFCVFTIGMFLQPLFQNTATTFTLHSDHSAFSSK"

TITLE

ORIGIN	AK053497	1864 bp	mrna	linear	HTC 20-SEP-2003
Query Match	18.7%;	Score 281.6;	DB 11;	Length 4046;	
Best Local Similarity	57.7%;	Pred. No. 2.5e-59;			
Matches 547;	Conservative 0;	Mismatches 389;	Indels 12;	Gaps 2;	
QY	99	CTACACACCTCGAGCGACCGGTGGCCACACGAGAGCGAACCCGCTAGAGTCTAGGTTGCG	158		
DB	331	CTATAACAAGCTGATCCGTCAGCTACTAATGGCTCTGAGCTGGTGAAGTCTGACAGTCTCAT	390		
QY	159	CTTGACCTTCGACGAATCATTTACGTGGACGAGAGAGATCAACTACTTATACCAATAT	218		
DB	391	GGTATCATTTGGCAGAGCTCATGTGTGACGCGGAGCAGATCATGACCAACCAAGT	450		
QY	219	ATGGCTGTCTGTGGAGTGAATGACTACAACCTGAGGTGGAACACAGCAGCAGTATGGCGG	278		
DB	451	CTGGCTGACCCAGAGTGGAGAGTATCGCTCTCATGGAAGCTGAGGATTCGACAA	510		
QY	279	GGTCAAGGACCTCAGATCAAGCCCAACAAGTTGTGGAACCGGAGCTCTTATGATATAA	338		
DB	511	TATGAAGAAAGTCCGACTCCCTTTAAGCACATCTGGCTCCAGATGTGTTCTATACAA	570		
QY	339	TAGTGTGACGAGGGTTTTCACGGACCTACAGACCAAGTGGTGGTGAAGCGCGG	398		
DB	571	CAATGCTGAGCGCATGTACGAAGTCTCTTCTATTTCCCAATGCTGTGGTCTCCCTATGATGG	630		
QY	399	CAGTTGCTGTAGTGTCCACTGGCACTTATTCAGAGCACATGCAAGATGACATCGCGTG	458		
DB	631	CAGCATCTTTTGCTACCGCTCCCTCTACAGAGCGCATGCAAGTTGAGGTGAAGCA	690		
QY	459	GTTCCTCTCGACGACCAACACTGTGATATGAAGTTGGTGGTGGACATATGACGGCAA	518		
DB	691	CTTCCCATTTGACGACGAGAACTGCACATGAAGTTCCGCTCTCGACCTACGACCGCAC	750		
QY	519	TCAGTTGGATCTGGTGTAAAGATGAGCGAGCGGGCGATCTATCGACTTCTATACAAA	578		
DB	751	TGAGATTGACTGGTGTCTAAAGCGCATGTGGCCAGC---CTGAGCACTTCACACCTAG	807		
QY	579	TGGGAGTGTGTATCTAATAGGAATGCCAGGCAAAAGAACACATACATACATCGGTGCTG	638		
DB	808	TGGGAGTGGACATCATCGCATGCGAGCCGACGACGAGAACCCAGACGACTCC--	865		
QY	639	CCCCGAGCCTCAGTGGACGTACCTTCCATCATCATGATGAAGAGCAACCTTTGACTA	698		
DB	866	-----ACCTAGCTGGACATCACCTACGACTTCATCATTCGTGCGAAACCGCTCTCTA	918		
QY	699	CTTCTTCAACTCATGTCCTCGGCGGTGCTGATCTCATCGATGCGCTCCTCGCTTCAC	758		
DB	919	CACATCAACTCATATCCCTCGGTACTCATCCTCGTGGCCATCCTGCTCTTCTA	978		
QY	759	ACTGCCACCACTCCGAGAGAACTCACACTTGGAGTCACTATTCTTCTATCGCTGAC	818		
DB	979	CCTGCCCTCAGACTGGTGGAAAGATGACACTTTGTATTCTGTGCTGTGGGCTCAC	1038		
QY	819	GGTGTCTCTCAACTGGTAGCGAGACCTTGCACAGGTTCTCGACGCTATCCCCCTGTT	878		
DB	1039	GGTGTCTCTCTCTCATCTCCAAAGATTGTCCTCCCACTCCCTCGACGTACGGCTGGT	1098		
QY	879	AGGAGCTACTTCAATTGCATCATGTTTCATGGTAGGCTGCTCTGTGTACTGACTGTGGT	938		
DB	1099	GGGAAGTACTCATGTTTCACTGTTGCTAGTCACTCTTCTCATGCTGCTAGCGGTG	1158		
QY	939	GGTACTCAATTACCACTGAAAGCTGATATACATGAATGCCAGTGGGATAAAATC	998		
DB	1159	TGTGCTCAATGTGCACACCGTTTCCGCTACCCAGCACACCATCGCGCCCTGGGTCAAGT	1218		
QY	999	AGTATTTCTTCAATGGTTGGATGATGACTGCGAATGTGAGGCGCAGG	1046		
DB	1219	GGTCTTCTCGGAAAGTGTGCCACCTCTCTCTCTCTGTCGACGACGACG	1266		
LOCUS DEFINITION	Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:EL10103E14 product:NEURONAL NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT, full insert sequence.				
ACCESSION	AK053497.1 GI:26343494				
VERSION	HTC; CAP trapper.				
KEYWORDS	Mus musculus (house mouse)				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)				
AUTHORS	99279253				
JOURNAL	MEDLINE 10349636				
PUBMED	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)				
AUTHORS	20493574				
JOURNAL	MEDLINE 11042159				
PUBMED	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Kageguchi, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)				
AUTHORS	20530913				
JOURNAL	MEDLINE 11076861				
PUBMED	4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)				
AUTHORS	5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)				
JOURNAL	MEDLINE 12011000				
PUBMED	6 (bases 1 to 1864)				
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hata, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akaira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.				
JOURNAL	Direct Submission				
PUBMED	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)				
AUTHORS	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.				
COMMENT					

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,

	Query Match	18.3%;	Score 275.6;	DB 11;	Length 3126;
	Best Local Similarity	54.7%;	Pred. No. 7.2e-58;		
	Matches 570;	Conservative 0;	Mismatches 469;	Indels 3;	Gaps 1;
QY	3	GGCCCTATGTTGGCGCCTTGGCGCTGCTGGCTTTGCTGCCGCTATCGGAGCAAGGTCC	62		
Db	187	GCGC GCGCTGCCATGCTGATGCTGTGGTCTCATGCTGCTGCAGATGGCCAGCGCCTCGGA	246		
QY	63	TCACGAGAAGAGACTCCCTGAAACGGCTTGCTGGCGAACTACAAACCCCTGGAGCGACCGGT	122		
Db	247	AGCTGAGCAGCGCCTGTTCCAGTACCTGTTTGAAGATTACAAAGAGATCATCCGGCCGGT	306		
QY	123	GGCCAAAGAGAGCGAAACGGCTAGAGGTGAGGTTCCGGTTGACCTTGCAGCAAAATCATTGA	182		

Db 307 GGCTAACGTTCCATCTCTGTCATCATCCAGTTTGGTGTCTATGCTCAGCTGGTGA 366
 QY 183 CGTGGACGAGAAATCACTTATTAACCAATATATGGCTGTCTGTTGAGTGGGAATGA 242
 Db 367 GGTGGATGAATGAATCAACAGATCATGGAACCAACCTGTGGCTGAAGCAAAATCTGGAATGA 426
 QY 243 CTACAACTGAGTGGTGAACGACAGCAGTATGCGGGTCAAGGACCTCAGGATCAGGCC 302
 Db 427 CTAACAGCTGAATGAACCCCTCTGACTACCAAGGGTGGAGTTTATGCGAGTCCCTGC 486
 QY 303 CAAACAAGTTGTGGAAGCGGACCTCTTATATATATAGTGTGAGAGGGTTTGAAGG 362
 Db 487 AGAGAAAGATCTGGAACACAGATCTGCTTTTACAAACGCGGATGGGATTTCCAAGT 546
 QY 363 GACCTACGACCAACGTTGTGTGTCAGAACGGGCGGAGTGTGCTGTACGTGCCACTGG 422
 Db 547 GGATGACAAACCAAGCTCTACTCAAGTACACAGGAGAAAGTGAATCTTGGATCCCTCGGC 606
 QY 423 CATATTCAAGACACATGCAAGATGGACATCGGTGTTTCCCTTCGACGACCAACACTG 482
 Db 607 CATCTTAAAGCTCATGCAAAATCGATGACTACTTCCGTTTGACTACCAAACTG 666
 QY 483 TGATATGAAGTTCCGTAGCTGGACATATGACGGCAATCAGTTGATCTGTGTAAAGA 542
 Db 667 CACCATGAAGTTCCGCTCTGGTCTCTACGACAGGCAAAAGATCGACCTGGTCTCTCA---T 723
 QY 543 TGAGGACGGCGGATCTATCGGACTTCTAACAATGGGAGTGGTATCTAATAGGAAT 602
 Db 724 TGGCTCTCTCATGAACCTCAAGGACTATTTGGAAAGTGGCGAGTGGGCCATCATTAAGC 783
 QY 603 GCCAGGCAAAAGAACCAATAATACATACGCTGTGCGCCGAGCCCTACGTGGAGCTCAC 662
 Db 784 CCGGGCTCACAACATGAATCAAGTACACTGTGTGAGGAGATCTTACCAAGACATCAC 843
 QY 663 CTTACACCATCATGATAAGAGAGCAAACTTTGTAATCTTCTTAACTGATCTGTGCTG 722
 Db 844 GTACTCGCTATACATTCGCGCTCTGCTGCTTCTACACCATCAACCTCATCATCTCCGTG 903
 QY 723 GGTCTGATCTCATGTGGACCTCTCGGCTTACACATGCTCCACAGCTCCGAGAGAA 782
 Db 904 CCGCTCATCT 963
 QY 783 ACTCACACTTGGAGTCACTATTTCTTCTATCTGCTGACGCTGTCTCTCAACCTGTAGCGGA 842
 Db 964 GGTGACGCTGTGATCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1023
 QY 843 GACCTGCGACAGGTTCTCGGAGCTATCCCTCTGTAGGAGCTACTTCAATTGATCAT 902
 Db 1024 GACCATCCCTTCCACCTCACTGTGTATCCCTTGTATCGGGAGTACCTCTCTTCACTAT 1083
 QY 903 GTTTCATGTTAGGCTGCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 962
 Db 1084 GATTTTGTACCTTGTCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 1143
 QY 963 AGCTGATATACATGAATGCCAGTGGATGAATAAATCAGTATTTCTTCAATGTTGCCATG 1022
 Db 1144 TCCGACCAACACACAGATGCCACTTGGGTCAAGGCTGTGTTTGTGAACCTTCTCCCGAG 1203
 QY 1023 GATCTGCGAATGTCGAGGCA 1044
 Db 1204 GGTATGTTTATGACTAGGCCA 1225

RESULT 15
 AY402876 1442 bp DNA linear GSS 15-DEC-2003
 LOCUS Homo sapiens CHRN4 gene, VIRTUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AY402876
 VERSION AY402876.1 GI:39758859
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1442)
 Clark A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 Science 302 (5652), 1960-1963 (2003)
 14671302
 2 (bases 1 to 1442)
 Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 Location/Qualifiers
 source 1..1442
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 gene <1..>1442
 /genes="CHRN4"
 /locus_tag="HCL1370"
 ORIGIN
 Query Match 18.3%; Score 274.4; DB 29; Length 1442;
 Best Local Similarity 57.0%; Pred. No. 9.8e-58;
 Matches 547; Conservative 0; Mismatches 401; Indels 12; Gaps 2;
 QY 100 TACAACTTGGAGCGGACCGGTGGCCAAACGAGAGCGAAACGCTAGAGGTCAAGTTCGCG 159
 Db 63 TACAATAACCTGATCCGCCAGCCAGCCAGCTTCTCAGCTCATCTCTCATCAAGCTGCAG 122
 QY 160 TTGACCTTGCAGCAAAATCATTGAGTGGACGAGAGAAATCAACTTATTATACCAATATA 219
 Db 123 CTCTCCCTGGCCAGCTTATCAGCGTGAATGAGCGAGAGAGATCATGACCACCAATGTC 182
 QY 220 TGGCTGTGTTGGAGTGGAAATGATCAAACTGAGGTGGAAAGACAGAGAGTATGGCGG 279
 Db 183 TGGCTGAAACAGGAATGACTGTATTACCGCTGTATCTGGAAACGCTCCCGTACGAGGT 242
 QY 280 GTCAGGACCTCAGGATCAGCCCAACAGAGTTGTGGAAGCGGACGCTCTTATGTATAAT 339
 Db 243 GTGAACATCTCTGAGGATCCCTTGAAGGCGATCTGGTTGCTGATCTGCTTTTACAC 302
 QY 340 AGTCTGACGAGGTTTTCAGCGGACCTTACAGACCAACGTTGGTGGTTCAGAAAGCGCGC 399
 Db 303 AACGCCAGGACCTATGAGTGTCTCTTACACCAACTTGATAGTCCGGTCCAAACGCG 362
 QY 400 AGTTGCTGTAGTGGCCACCTGGGATATTCAAGAGCATGCAAGATGGAATCGCTGG 459
 Db 363 AGCGTCTGTGGCTGCCCTTGCCTCTTACAAAGAGCGCTCAAGATTGAGGTGAAGTAC 422
 QY 460 TTTCCCTTCGACGACCACTGTGATATGAATGAGTTCCGTAGCTGACATATGACGGCAAT 519
 Db 423 TTTCCCTTCGACGACCACTGTGACCTTCAAGTTCCGCTCTGAGACCTATGACCAACG 482
 QY 520 CAGTTGGATCTGGTCTAAAAAGATGAGGACGCGCGGATCTATCGGATCTCATTAACAAAT 579
 Db 483 GAGATAGACATGGTCTCATGACGCCCAACAGCCAGC---ATGGATGACTTTACTCCAGT 539
 QY 580 GGGGAGTGTATCTATATAGGAATGCCAGCAAAAGACACAATTAACATAGCGTGTGC 639
 Db 540 GGTGAGTGGACATAGTGGCCCTCCAGGGAGAGGACAGTGAACCCCAAGACCCAG- 598
 QY 640 CCCGAGCCCTACGTGGAGCTCACCTTCCACCATCATGATGAAGAGCAACCTTGTACTAC 699

```
Db 599 -----CTACGTGGACGTGACTTAGCACTTCATCATCAAGCGAAGCCTCTGTCTTCTAC 650
QY 700 TTCTTTCAACCTGATCGTCCGGTCGGTGTGATCTCATTCGATGGCACTCCTCGGCTTCACA 759
Db 651 ACCATCAACCTCATCATCCCTCGTGTGCTCACCACCTTGTGGCCATCCTCGTCTTCTAC 710
QY 760 CTGCCACCAAGACTCCGGAGAGAACTCACACTTGGAGTCACTATTCTTCTATCGCTGACG 819
Db 711 CTGCCATCCGACTCGCGCGAAGATGACACTGTGCATCTCAGTGTCTGGCACTGACA 770
QY 820 GTGTTCTCAACCTGGTAGCCGAGACCCTGCCACAGGTCTCCGACGCTATCCCCCTGTTA 879
Db 771 TTCTTCTGCTGCTCATCTCTCAAGATCGTGCCACCCACCTCCCTCGATGTGCTCTCATC 830
QY 880 GGGACGTACTTCAATTGCATCATGTTTCATGGTAGCGTCTGTGGTACTGACTGTGGTG 939
Db 831 GGCAAGTACCTCATGTTCAACCATGGTGTGCTGCACCTTCTCATCGTCAACGCGTCTGT 890
QY 940 GTACTCAATTACCACCATCGAACAGCTGATATATACATGAATGCCACAGTGGATAAATCA 999
Db 891 GTGCTCAATGTGACCAACCGCTCGCCAGCACCCACACCATGGCACCCCTGGGTCAAGCGC 950
QY 1000 GTATTCCTACAATGGTTGCCATGATGATGCGGAATGTCCGAGGCCAGGGAAGAGATCAC 1059
Db 951 TGGTTCTCTGCACAAAGCTGCTACCTTCTTCTTCAAGCGCCCTGGCCCCGACAGCAGC 1010
```

Search completed: May 8, 2004, 06:38:57
Job time : 4044.19 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:09:28 ; Search time 43.0968 Seconds
(without alignments)

3284.615 Million cell updates/sec

Title: US-09-303-232-6

Perfect score: 2640

Sequence: 1 MAPMLAALALLPVSEOG.....LFTIATVAVLLSAPHIIVQ 501

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lasting first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2640	100.0	501	3 AAY50816	Aay50816 H. viresc
2	1803.5	68.3	496	3 AAY50815	Aay50815 H. viresc
3	1609	60.9	770	3 AAY50814	Aay50814 D. melano
4	1570.5	59.5	498	4 ABB60432	Abb60432 Drosophil
5	1283	48.6	311	4 ABB63683	Abb63683 Drosophil
6	1258.5	47.7	502	2 AAW44153	Aaw44153 Human neu
7	1258.5	47.7	502	2 AAW09025	Aaw09025 Neuronal
8	1258.5	47.7	502	3 AAB24088	Aab24088 Human PRO
9	1258.5	47.7	502	4 AAB82690	Aab82690 Nicotinic
10	1258.5	47.7	502	4 AAB50012	Aab50012 Wild-type
11	1258.5	47.7	502	5 ABB70492	Abb70492 Human neu
12	1258.5	47.7	502	5 ABB82435	Abb82435 Human neu
13	1258.5	47.7	502	7 ADA10874	Ada10874 Human neu
14	1258.5	47.7	502	7 ADA47051	Ada47051 Human PRO
15	1258.5	47.7	502	7 ADE57310	Ade57310 Human PRO
16	1254.5	47.5	502	2 AAW69216	Aaw69216 V274T var
17	1253	47.5	502	2 AAW12368	Aaw12368 Neuronal
18	1252.5	47.4	502	4 AAB50015	Aab50015 Mutant hu
19	1248.5	47.3	502	4 AAB50016	Aab50016 Mutant hu
20	1246.5	47.2	502	7 AAB50017	Aab50017 Rat Prote
21	1246.5	47.2	502	7 AAD47049	Ada47049 Rat Prote
22	1242.5	47.1	502	4 AAB50017	Aab50017 Mutant hu
23	1226.5	46.5	511	2 AAW12369	Aaw12369 Neuronal
24	1080	40.9	554	4 AAE12824	Aae12824 Caenorhab
25	1080	40.9	554	6 ABP96318	Abp96318 Caenorhab

ALIGNMENTS

RESULT 1

AAAY50816
ID AAY50816 standard; protein; 501 AA.

XX AC AAY50816;

XX DT 17-FEB-2000 (first entry)

XX DE H. virescens acetyl-choline receptor protein from clone Hva7-2.

XX KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
XX NW neurotransmission; plant protection agent; conductance; AChR.

XX OS Heliothis virescens.

XX PN DE19819829-A1.

XX PD 11-NOV-1999.

XX PF 04-MAY-1998; 98DE-01019829.

XX PR 04-MAY-1998; 98DE-01019829.

XX (FARB) BAYER AG.

XX PI Adamczewski M, Oellers N, Schulte T;

XX DR WPI; 2000-014207/02.

XX N-PSDB; AAZ24477.

XX PT New nucleic acid encoding a nicotinic acetylcholine receptor from

XX insects, used to identify potential insecticides.

XX Example 1a; Page 22-23; 26pp; German.

XX CC This invention describes a novel nucleic acid (NA) encoding a nicotinic
XX acetyl-choline receptor (I) from insects which can be used as an
XX insecticide. Inhibitors of (I) interfere with neurotransmission. (I)
XX (also vectors containing it, its regulatory regions, and antibodies
XX directed against (I)-encoded proteins) are used to screen for: (a) plant
XX protection agents that alter conductance of AChR, potentially useful as
XX insecticides, or (b) genes which encode polypeptides that are involved in
XX formation of functionally related AChR in insects. (I) are also used to
XX isolate and characterize the specified regulatory regions and for
XX recombinant production of (II). This sequence represents an acetyl-
XX choline receptor isolated from *Heliothis virescens*
XX Sequence 501 AA;

Ades7314 Human Pro
Ade57318 Human Pro
Aaw09022 Neuronal
Ada10857 Human neu
Ada83810 Human CHR
Ade57312 Rat Prote
Ade57316 Rat Prote
Aab50014 ChimERIC
Abb82431 Human neu
Abb61954 Drosophil
Aab50018 Mature ce
Aaw44156 Human neu
Aaw09018 Neuronal
Abb70491 Human neu
Abb82434 Human neu
Ada10863 Human neu
Abb62727 Drosophil
Aaw44155 Human neu
Aar73966 Alpha 2 s

26 989.5 37.5 503 7 ADE57314
27 989.5 37.5 503 7 ADE57318
28 989.5 37.5 504 2 AAW09022
29 989.5 37.5 504 7 ADA10857
30 982.5 37.2 502 6 ADA83810
31 977.5 37.0 499 7 ADE57312
32 977.5 37.0 499 7 ADE57316
33 962.5 36.5 470 4 AAB50014
34 946.5 35.9 504 5 ABB70488
35 946.5 35.9 504 5 ABB82431
36 943 35.7 576 4 ABB61954
37 942 35.7 448 4 AAB50018
38 940.5 35.6 504 2 AAW44156
39 933 35.3 494 2 AAW09018
40 933 35.3 494 5 ABB70491
41 933 35.3 494 5 ABB82434
42 933 35.3 494 7 ADA10863
43 933 35.3 580 4 ABB62727
44 922.5 34.9 529 2 AAW44155
45 922.5 34.9 529 2 AAR73966

Query Match 100.0%; Score 2640; DB 3; Length 501;
 Best Local Similarity 100.0%; Pred. No. 1.2e-267;
 Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPMLAALALLPVSEQGPHEKRLINALLANVTLEPVAESEPLEVRFGLTQQII 60
 DB 1 MAPMLAALALLPVSEQGPHEKRLINALLANVTLEPVAESEPLEVRFGLTQQII 60

QY 61 DVEKNQLITNTWLSLEWNYLRNDSYGGVKDLRITPNKLPDVLVMSADRGFD 120
 DB 61 DVEKNQLITNTWLSLEWNYLRNDSYGGVKDLRITPNKLPDVLVMSADRGFD 120

QY 121 GTYQTNVVRSGGSLVPPGIFKSTCKMDIANFPDDQDCHDKMFGSWTYDGNQLDLVK 180
 DB 121 GTYQTNVVRSGGSLVPPGIFKSTCKMDIANFPDDQDCHDKMFGSWTYDGNQLDLVK 180

QY 181 DEAGGDSDFITNGEWYLIQMPGKNTITYACCEPYVDVTFIMIRRTLYFFNLIVP 240
 DB 181 DEAGGDSDFITNGEWYLIQMPGKNTITYACCEPYVDVTFIMIRRTLYFFNLIVP 240

QY 241 CVLISSMALLGFTLPDPSGKLTGLVTLISLTVFLNLVAETLPQVSDAIPLLGTYFNCI 300
 DB 241 CVLISSMALLGFTLPDPSGKLTGLVTLISLTVFLNLVAETLPQVSDAIPLLGTYFNCI 300

QY 301 MFVASSVLTVVVLYNHHRTADIHMPQWIKSVFLQWLPWILRMSRPGKTKITIMN 360
 DB 301 MFVASSVLTVVVLYNHHRTADIHMPQWIKSVFLQWLPWILRMSRPGKTKITIMN 360

QY 361 TRMRELEKERSKSLANVLDDDDFRHGGPPPNSTASTGNLPGGCSIFRTDPRRSFVR 420
 DB 361 TRMRELEKERSKSLANVLDDDDFRHGGPPPNSTASTGNLPGGCSIFRTDPRRSFVR 420

QY 421 PSTMEDVGGGLGSHHRELHLILRELQFITARMKKADEAEELISDWKFAAMVVDRCFLVF 480
 DB 421 PSTMEDVGGGLGSHHRELHLILRELQFITARMKKADEAEELISDWKFAAMVVDRCFLVF 480

QY 481 TLFITIAVALLSAPHIIQV 501
 DB 481 TLFITIAVALLSAPHIIQV 501

RESULT 2
 AAY50815
 ID AAY50815 standard; protein; 496 AA.
 AC AAY50815;
 DT 17-FEB-2000 (first entry)
 XX H. virescens acetyl-choline receptor protein from clone Hva7-1.
 DE Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
 XX neurotransmission; plant protection agent; conductance; ACHR.
 OS Heliothis virescens.
 XX DE19819829-Al.
 XX 11-NOV-1999.
 XX 04-MAY-1998; 98DE-01019829.
 XX 04-MAY-1998; 98DE-01019829.
 XX (FARB) BAYER AG.
 XX Adamczewski M, Oellers N, Schulte T;
 XX WPI; 2000-014207/02.
 XX N-PSDB; AAZ24476.
 XX New nucleic acid encoding a nicotinic acetylcholine receptor from

PT insects, used to identify potential insecticides.
 PS Example 1a; Page 17-19; 26pp; German.
 XX This invention describes a novel nucleic acid (NA) encoding a nicotinic
 CC acetyl-choline receptor (I) from insects which can be used as an
 CC insecticide. Inhibitors of (I) interfere with neurotransmission. (I)
 CC (also vectors containing it, its regulatory regions, and antibodies
 CC directed against (I)-encoded proteins) are used to screen for: (a) plant
 CC protection agents that alter conductance of ACHR, potentially useful as
 CC insecticides, or (b) genes which encode polypeptides that are involved in
 CC formation of functionally related ACHR in insects. (I) are also used to
 CC isolate and characterize the specified regulatory regions and for
 CC recombinant production of (II). This sequence represents an acetyl-
 CC choline receptor isolated from *Heliothis virescens*
 XX Sequence 496 AA;

Query Match 68.3%; Score 1803.5; DB 3; Length 496;
 Best Local Similarity 68.7%; Pred. No. 8.2e-180;
 Matches 347; Conservative 53; Mismatches 78; Indels 27; Gaps 7;

QY 2 APMLAALALLPVSEQGPHEKRLINALLANVTLEPVAESEPLEVRFGLTQQII 60
 DB 12 APAGLLLLCLLWPRGARCGYHEKRLIHLLDHYNVLERPVVNESDPLQLSFGLTQQII 71

QY 61 DVEKNQLITNTWLSLEWNYLRNDSYGGVKDLRITPNKLPDVLVMSADRGFD 120
 DB 72 DVEKNQLITNTWLSLEWNYLRNDSYGGVKDLRITPNKLPDVLVMSADRGFD 131

QY 121 GTYQTNVVRSGGSLVPPGIFKSTCKMDIANFPDDQDCHDKMFGSWTYDGNQLDLVK 180
 DB 132 STYPTNVVRNNGSLVPPGIFKSTCKIDITWFPDDQDCEMKFGSWTYDGNQLDLVK 191

QY 181 DEAGGDSDFITNGEWYLIQMPGKNTITYACCEPYVDVTFIMIRRTLYFFNLIVP 240
 DB 192 DEGGGDISFVINGEWELIGVPGKNEIYVNCCEPYDITFAVIRRTLYFFNLIVP 251

QY 241 CVLISSMALLGFTLPDPSGKLTGLVTLISLTVFLNLVAETLPQVSDAIPLLGTYFNCI 300
 DB 252 CVLISSMALLGFTLPDPSGKLTGLVTLISLTVFLNLVAETLPQVSDAIPLLGTYFNCI 311

QY 301 MFVASSVLTVVVLYNHHRTADIHMPQWIKSVFLQWLPWILRMSRPGKTKITIMN 360
 DB 312 MFVASSVLTVVVLYNHHRTADIHMPQWIKSVFLQWLPWILRMSRPGSATPPARV- 370

QY 361 TRMRELEKERSKSLANVLDDDDFRHGGPPPNSTASTGNLPGGCSIFRTDPRRSFVR 420
 DB 371 PPPDLELRERSKSLANVLDDDDFRH- PQAQQ-----PQCRY---YRGG--- 414

QY 421 PSTMEDVGGGLGSH- HRELHLILRELQFITARMKKADEAEELISDWKFAAMVVDRCF 475
 DB 415 -----EENGAGLAHSCFGVDYELSLILKEIRVITDQMKRDEDEDADISRDWKFAMVVDRL 470

QY 476 CLFVFTLFTIATVALLSAPHIIV 500
 DB 471 CLFIITLFTIATVALLSAPHIMV 495

RESULT 3
 AAY50814
 ID AAY50814 standard; protein; 770 AA.
 XX AAY50814;
 AC AAY50814;
 XX 17-FEB-2000 (first entry)
 DT D. melanogaster acetyl-choline receptor protein from clone Da7.
 DE Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
 XX neurotransmission; plant protection agent; conductance; ACHR.
 XX Drosophila melanogaster.

XX DE19819829-A1.
PN 11-NOV-1999.
XX 04-MAY-1998; 98DE-01019829.
XX 04-MAY-1998; 98DE-01019829.
PR (FARB) BAYER AG.
XX Adamczewski M, Oellers N, Schulte T;
XX WPI: 2000-014207/02.
DR N-PSDB; AAZ24475.
XX New nucleic acid encoding a nicotinic acetylcholine receptor from
PT insects, used to identify potential insecticides.
PS Example 1a; Page 12-14; 26pp; German.
XX This invention describes a novel nucleic acid (NA) encoding a nicotinic
CC acetyl-choline receptor (I) from insects which can be used as an
CC insecticide. Inhibitors of (I) interfere with neurotransmission. (II)
CC (also vectors containing it, its regulatory regions, and antibodies
CC directed against (I)-encoded proteins) are used to screen for: (a) plant
CC protection agents that alter conductance of AChR, potentially useful as
CC insecticides, or (b) genes which encode polypeptides that are involved in
CC formation of functionally related AChR in insects. (I) are also used to
CC isolate and characterize the specified regulatory regions and for
CC recombinant production of (II). This sequence represents an acetyl-
CC choline receptor isolated from *Drosophila melanogaster*
XX
XX Sequence 770 AA;
Query Match 60.9%; Score 1609; DB 3; Length 770;
Best Local Similarity 60.8%; Pred. No. 4.2e-159;
Matches 319; Conservative 57; Mismatches 65; Indels 84; Gaps 10;
QY 10 LLALLPVSEQ-----GPEKRLNALLANTYTLRVPVANESEPLEVFGTLTQIIVDEK 65
DB 295 LITLNLAKVCLAGYHEKRLHLDLPYTLRPNVNESDPLQLSFGTLTQIIVDEK 354
QY 66 NOLLITNLSLEWNDYLRWDSYGGVKDLRITPKLWKPDLVLYNSADGFGTYQT 125
DB 355 NQLLVTVNWLKLEWNNLWNTSDTGGVKDLRIPHRLWKPDLVLYNSADGFGTYQT 414
QY 126 NVVVRSGSCLYVPPGIFKSTCKMDIAWFPFDDQHCMDKFGSWTYDGNQLDLVKDEAG 185
DB 415 NVVVRNNGSCLYVPPGIFKSTCKIDITWFPFDDQRCMEKFGSWTYDGFQDLQLQDETGG 474
QY 186 DLSDFITNGEWLYGMGPKNITYIACCPBPYVDVTFIMIRRTLYFFNLIIVCVLIS 245
DB 475 DISSVYLVNGEWELLYGPGKRNELIYNCOPEYIDITFAIIRRTLYFFNLIIVCVLIA 534
QY 246 SMALLGFTLPPSGEKLTLGVTLISLTVFLNLVAETLPQVSDAIPLLGTFFNCFMFWA 305
DB 535 SMALLGFTLPPSGEKLTLGVTLISLTVFLNLVAETMPATSDAVPL----- 581
QY 306 SSVLTVVLYNHHRTADIHMPQWTKSVFLOLWPLWILMRSPGKKITRK---TIMNTR 362
DB 582 -----WIRVFLCWLPWILMRSPGRLLILEFFTPCSDTS 617
QY 363 -----MRELEKERSKSLANVLIDDDDFRHG---PPPNSTAGTGMGPGCSIFRTD 413
DB 618 SERKHQILSDVELKERSKSLANVLIDDDDFRHNCRPMTGP-----GTLPHNPAFYRTV 672
QY 414 FRRSFVRPSTMDVVG--GGIGS-----HHR-----ELHLTLRELQFTARMKA 455
DB 673 YCGG-----DDGIGIPGSTRMPDVAHTTCIKSSTSEYELGLTLKIRFITDQLRKD 724
QY 456 DEEAELISDKFAAMVVDRCFLVFVFTLTIIATVAVLLSAPHIIV 500
|| : : : ||||| || : : : ||||| || : : : ||||| || : : : ||||| |||||

DB 725 DECNDIANDWKFAAMVVDRLCLIFITMFAILLATIAVLISAPHIIV 769
RESULT 4
ID ABB60432
XX ABB60432 standard; protein; 498 AA.
AC ABB60432;
XX 26-MAR-2002 (first entry)
XX *Drosophila melanogaster* polypeptide SEQ ID NO 8088.
DE *Drosophila*; developmental biology; cell signalling; insecticide;
XX *Drosophila*; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX *Drosophila melanogaster*.
OS WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI: 2001-656860/75.
DR N-PSDB; ABL04535.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
PT interactions.
XX Disclosure; SEQ ID NO 8088; 2lpp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 498 AA;
Query Match 59.5%; Score 1570.5; DB 4; Length 498;
Best Local Similarity 61.1%; Pred. No. 2.4e-155;
Matches 319; Conservative 44; Mismatches 84; Indels 75; Gaps 9;
QY 2 APMLAALAL-----LALLPVSEQGPEKRLNALLANTYTLRVPVANESEPLEVFGTL 55
DB 3 SPLPASLSLFLVLLFLAIKESQCPGPEKRLNALLSTYTLRVPVANESEPLEVFGTL 62
QY 56 LQIIDVDKKNOLLITNLSLEWNDYLRWDSYGGVKDLRITPKLWKPDLVLYNSA 115
DB 63 LQIIDVDKKNOLLITNLSLEWNDYLRWDSYGGVKDLRITPKLWKPDLVLYNSA 122
QY 116 DSGFGDTGYTNVVRSGSCLYVPPGIFKSTCKMDIAWFPFDDQHCMDKFGSWTYDGNOL 175
DB 123 DSGFGDTGYTNVVRSGSCLYVPPGIFKSTCKIDITWFPFDDQHCMDKFGSWTYDGNOL 182
QY 176 DLVLKDEAGGDLSDFITNGEWLYGMGPKNITYIACCP-----EPYVDVT 221
DB 183 DLVLNSEDGDLSDFITNGEWLYGMGPKNITYIACCP-----EPYVDVT 239
|| : : : ||||| || : : : ||||| || : : : ||||| || : : : ||||| |||||

```
QY 222 FTIMIRRRTYFFNLIIVPCVLSSMALL-----GFTLPDPSGKLTGLV----- 266
Db 240 -----ILFPFNCAMCANLIDGPTGLHIAAGFGRETDAGNYTTIINSIKPCRHH 291
QY 267 -----TILSLTVFLNVAETLPQVSDAIPLLGTGFYFNCIMFMWASSVVLTV 312
Db 292 ADVNGCCSSRYRTHILKRYSLKYPGPIGTIDHYQC-RAESTVFCIMFMWASSVVLTV 350
QY 313 VVLNYHRTADIHEMPWIKSVFLOLWILRMSRPGKTKTRTIMNTMRELELKERS 372
Db 351 VVLNYHRTADIHEMPWIKSVFLOLWILRMSRPGKTKTRTIMNTMRELELKERS 410
QY 373 SKLLANVLDDIDDFRHGPPPPNSTAGTGNLGPCCSIFRTDFRSFVRPSTMEDVGGGLG 432
Db 411 SKLLANVLDDIDDFRH-----TISGQTAIGSS-----ASFGRTTVEHHTAIG 456
QY 433 SHRELHLILRELQFITARMKKADEAEALISDMKFAAMVVDR 474
Db 457 CNEKHLHLILKELQFITARMKKADEAEALIGDMKFAAMVVDR 498

RESULT 5
ABB63683
ID ABB63683 standard; protein; 311 AA.
XX
AC ABB63683;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 17841.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
PI WPI; 2001-656860/75.
XX
DR N-PSDB; ABL07786.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 17841; 2lpp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 311 AA;
```

Query Match 48.6%; Score 1283; DB 4; Length 311;
Best Local Similarity 79.4%; Pred. No. 1.8e-125;

```
Matches 231; Conservative 30; Mismatches 30; Indels 0; Gaps 0;
QY 4 MLAALALLLPVSEOGPEKRLINALLANYNTERPVANESPELVRFGLTQQIITDVD 63
Db 20 LVTYGLGLINIPACAGPHEKRLHLLHALLDNYSLERPVNVNESDPLQSLFGLTLMQIITDVD 79
QY 64 EKNQLLITNWLSEWNDYNLRWDSYGGVKDLRITPNKLWKPDLVLYNSADEGFGDTY 123
Db 80 EKNQLLITNWLKLEWNDYNLRWNSSEFGVGRDLRIPPHRLWKPDLVLYNSADEGFGDTY 139
QY 124 QTNVVRSGSCLYVPGPIFKSTCKMDIAWFPDDQHCMDKFGSWTYDGNQLDLVLKDEA 183
Db 140 ATNVVNRNGSCLYVPGPIFKSTCKDITWFPDDQRCCKMFGSWTYDGFQDLQLQDEA 199
QY 184 GGLSDFITNGEWWLGMPCGKNTIYACCPPEYVDVTFIMIRRRTYFFNLIIVPCVL 243
Db 200 GSDISSFITNGEWDLLGVPGKRNEIYNNCCPEYIDITFAILLRRKTYFFNLIIVPCVL 259
QY 244 ISSMALLGFTLPDPSGKLTGLVTILSLTVFLNVAETLPQVSDAIPLLG 294
Db 260 IASMALLGFTLPDPSGKLSGLVTILSLTVFLNVAETMPATSDAIPLLG 310

RESULT 6
AAW44153
ID AAW44153 standard; protein; 502 AA.
XX
AC AAW44153;
XX
DT 14-MAY-1998 (first entry)
XX
DE Human neuronal nicotinic acetylcholine receptor alpha-7 subunit.
XX
KW Human; neuronal nicotinic acetylcholine receptor; alpha-7 subunit;
KW brain tissue; screening; NACHR; antibody.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT Peptide 1..23
FT Domain /label= signal
FT Domain 229..256
FT Domain /label= TMD1
FT Domain /note= "transmembrane domain"
FT Domain 262..284
FT Domain /label= TMD2
FT Domain /note= "transmembrane domain"
FT Domain 290..317
FT Domain /label= TMD3
FT Domain /note= "transmembrane domain"
FT Domain 318..461
FT Domain /label= cytoplasmic_loop
FT Misc-difference 343 /note= "encoded by CTS"
FT Domain 462..487
FT Domain /label= TMD4
FT Domain /note= "transmembrane domain"
XX
PN WO9420617-A2.
XX
PD 15-SEP-1994.
XX
PF 08-MAR-1994; 94WO-US002447.
XX
PR 08-MAR-1993; 93US-00028031.
XX
PA (SALK ) SALK INST BIOTECHNOLOGY IND ASSOC.
XX
PI Elliott KJ, Ellis SB, Harpold MM;
XX
PI WPI; 1994-303024/37.
XX
DR N-PSDB; AAV12197.
```


Db 461 KFAACVVDRLCLMAFSVFTIICITIGILMSAPNFV 494

RESULT 8

ID AAB24088 standard; protein; 502 AA.

XX AC AAB24088;

XX DT 29-JAN-2001 (first entry)

XX DE Human PRO2145 protein sequence SEQ ID NO:77.

XX KW Human; tumour; diagnosis; neoplastic disease; neoplastic cell growth; proliferation; tumorigenesis; identification; cancer; cytostatic; neurotropic; neuroprotective; antiinflammatory; immunosuppressive; immunostimulant; antiangiogenic; leukaemia; lymphoid malignancy; neuronal disorder; glioma disorder; astrocytic disorder; angiogenic; hypothalamic disorder; glandular disorder; macrophagal disorder; epithelial disorder; stromal disorder; blastocoelec disorder; inflammatory disorder; immunologic disorder.

XX OS Homo sapiens.

XX KW WO200053755-A2.

XX PN 14-SEP-2000.

XX PD 06-JAN-2000; 2000WO-US000376.

XX PF 08-MAR-1999; 99WO-US005028.

XX PR 02-JUN-1999; 99WO-US012252.

XX PR 23-JUN-1999; 99US-0141037P.

XX PR 07-JUL-1999; 99US-0143048P.

XX PR 26-JUL-1999; 99US-0145698P.

XX PR 30-NOV-1999; 99WO-US028313.

XX PR 20-DEC-1999; 99WO-US030911.

XX PR 05-JAN-2000; 2000WO-US000219.

XX PA (GETH) GENENTECH INC.

XX PI Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hillan KJ, Roy MA; Watanabe CK, Wood WI;

XX DR WPI; 2000-572270/53.

XX DR N-PSDB; AAC58395.

XX PT Thirty PRO polynucleotides encoding PRO polypeptides, useful in the treatment, diagnosis and prevention of cancer.

XX PS Claim 61; Fig 58; 286pp; English.

XX CC The present invention describes an isolated antibody that binds to one of the human PRO proteins designated PRO212, PRO290, PRO341, PRO355, PRO619, PRO717, PRO809, PRO830, PRO848, PRO943, PRO1005, PRO1009, PRO1025, PRO1030, PRO1097, PRO1107, PRO1111, PRO1153, PRO1182, PRO1184, PRO1187, PRO1281, PRO23, PRO39, PRO834, PRO1317, PRO1710, PRO2094, PRO2145 OR PRO2198. PRO antagonists can be used to inhibit tumour cell growth. The PRO polypeptides and nucleotides are useful in the treatment, diagnosis and prevention of cancer. The antibodies and other anti-tumour compounds may be used to treat various conditions, including those characterised by overexpression and/or activation of the amplified PRO genes. Exemplary conditions or disorders to be treated with such antibodies and other compounds include benign or malignant tumours (e.g., renal, liver, kidney, bladder, breast, gastric, ovarian, colorectal, prostate, pancreatic, lung, vulva, thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head and neck tumours), leukaemias and lymphoid malignancies, other disorders such as neuronal, glial, astrocytic, hypothalamic and other glandular, macrophagal, epithelial, stromal and blastocoelec disorders, and inflammatory, angiogenic and immunologic disorders. AAC58242 to AAC58366 represent PCR primers and hybridisation probes used in the isolation of the human PRO sequences. AAC58367 to AAC58396 and AAB24057 to AAB24089 represent human PRO

CC polynucleotide and protein sequences given in the exemplification of the
CC present invention
XX Sequence 502 AA;

Query Match 47.7%; Score 1258.5; DB 3; Length 502;
Best Local Similarity 48.6%; Pred. No. 1.4e-122;
Matches 250; Conservative 75; Mismatches 138; Indels 51; Gaps 7;

Qy 8 LALLA-LLPVSEOGPEHEKRLINALLANYTLERLVANESPLEVRFGLTQQIIDVDEKN 66
Db 10 LALAASLLHVSLOGEFQKLYKELVKYNPLERPVANDSQPLTFYFSLSLQIINDVDEKN 69
Qy 67 QLLITNIWLSEWINDYLRWDSBYGGVKDLRITPNKLPKDPVLMVNSADEGFGTQTN 126
Db 70 QVLTNIWLQMSWTDHYLQWNVSEYPGKTVRPDQGIWKPDILLVNSADERFATHTN 129
Qy 127 VVVRSGSCLYVPGIFKSKCKMDIAWFPDDQHCMDKMGFSWTYDGNLDLVLKDEAGD 186
Db 130 VLVNSSGHCQYLPPIGFKSSCYIDVRWFFPDVQCKLKFGSWSYGGMSLDLQMQE---AD 186
Qy 187 LSDFITNGEWYLLGMPGKNITTYACCPYVDVVTFTIMIRRTLYYFFNLIVPCVLIS 246
Db 187 ISGYIPNGEWDLVGIPGKRSERFECCKEYPDVFTVTWRRRTLYYGLNLLIECVLISA 246
Qy 247 MALLGFTLPDPSGKLTGVTILLSLTVFLNLVAETLPQVSDAIPILGTFFNCFMFMVAS 306
Db 247 LALLVFLPADSGEKISLGITVLLSLTVFMLVAEIMPATSDSVFLTAQIFASTIIVGL 306
Qy 307 SVLVTVVLNHYHRTADIEHMPQWIKSVFLQWLPWILMRSPGKKITRKTIMMTRREL 366
Db 307 SVVTVIVLQVHHDDPDGKMKPWTRVILLNWCWFLMRKRPBGDKVRPAQCHKQRCSL 366
Qy 367 ELKERSKSLANVLDDDDFRHGGPPPNSTASGNL-----GPGC 407
Db 367 ASVMSAVA-----PPP-----ASGNLLYIGFRLGDGVHCVPTPDSGVVC 407
Qy 408 SIFRTDFRSFVRPSTMEDV--GGGLGSHHREHLHLRELQFITARMKKADEAEELISDW 465
Db 408 G-----RMACSPTHDEHLHGQPEGDPDLAKILLEEVRYIANRRCQDESAVCSW 460
Qy 466 KFAAMVDRPCLFVFTLTIIATVAVLLSAPHII 499
Db 461 KFAACVVDRLCLMAFSVFTIICITIGILMSAPNFV 494

RESULT 9

ID AAB82690 standard; protein; 502 AA.

XX AC AAB82690;

XX DT 15-OCT-2001 (first entry)

XX DE Nicotinic acetylcholine receptor alpha7.

XX KW Nicotinic acetylcholine receptor; nAChR; human;
KW acetylcholine binding protein; AChBP; mollusc; ligand-binding protein;
KW ligand-gated ion channel; crystal; drug design; protein co-ordinate data;
KW schizophrenia; Alzheimer's disease; nicotine addiction;
KW Tourette's syndrome; therapy; neurotropic; neuroprotective.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers
FT Region 108..115

FT /note= "conserved ligand-binding region, residues Trp108 and Tyr115 are essential"

FT Region 171..173

FT /note= "conserved ligand-binding region, residues Trp171 and Tyr173 are essential"

FT Region 210..217

FT /note= "conserved ligand-binding region, residues Tyr210,

Cys212, Cys213 and Tyr217 are essential"

PT XX WO200158951-A2.
 XX XX 16-AUG-2001.
 XX PF 09-FEB-2001; 2001WO-EP001457.
 XX XX 10-FEB-2000; 2000EP-00200443.
 XX PR 31-OCT-2000; 2000EP-00203810.
 XX XX (TEWE-) STICHTING TECH WEIENSCHAPPEN.
 XX PA Smit AB, Sixma TK;
 XX PI WPI; 2001-497071/54.
 XX DR
 XX XX Water-soluble ligand-binding proteins derived from molluscs and analogs
 PT of ligand-gated ion channels, useful in drug screening assay, where the
 PT drugs identified can be used in the treatment of Alzheimer's disease or
 PT schizophrenia.
 XX PS Disclosure; Page 252-254; 260pp; English.
 XX CC The present sequence is that of the alpha subunit of human nicotinic
 CC acetylcholine receptor (nAChR). The sequence includes regions that are
 CC conserved throughout the various nAChR alpha subunits and which are
 CC essential for ligand binding. The invention relates to water-soluble
 CC ligand-binding proteins derived from molluscs, especially acetylcholine-
 CC binding proteins (AChBPs) and analogues of ligand-gated ion channels,
 CC their crystals, and their use for screening ligands of ligand-gated ion
 CC channels. The water-soluble ligand-binding proteins are capable of
 CC forming multimers and are amenable to crystallization. The crystal
 CC structure of AChBP is provided, and can be used to generate 3D models of
 CC the extracellular ligand-binding domain of ligand-gated ion channels and
 CC thus for screening of drugs that act on these ion channels. Chimeric
 CC proteins are provided that are capable of binding a ligand of a ligand-
 CC gated receptor, and comprise at least the amino acids of the AChBP
 CC determining solubility of the AChBP, in the same positions as in the
 CC AChBP, and also comprising amino acids determining binding to the ligand.
 CC In the chimeric proteins, at least the essential amino acids of at least
 CC 1 of the conserved regions of an nAChR have been substituted for the
 CC corresponding amino acids, and preferably entire stretches have been
 CC substituted. New drugs can be developed that selectively intervene in
 CC neuronal signalling pathways, especially where the ligand-gated ion
 CC channel is the nAChR, and the related disorder is Tourette's syndrome,
 CC Alzheimer's disease, addiction to nicotine or schizophrenia
 XX SQ Sequence 502 AA;
 Query Match 47.7%; Score 1258.5; DB 4; Length 502;
 Best Local Similarity 48.6%; Pred. No. 1.4e-122;
 Matches 250; Conservative 75; Mismatches 138; Indels 51; Gaps 7;
 QY 8 LALLA-ILPVSEOGHEKRLNALLANVTLPVPANSEPLEVFGTLQOIIVDDEKN 66
 DB 10 LALAASLLHVSLSQGEFQRKLYKELYNPLRPVANDSQPLTVYFVLSLQIMVDDEKN 69
 QY 67 QLLITNIWLSLEWNYDNLRWNDSEYGVKDLRITPNKLWKPDLVLMYNSADEGFGDTYQTN 126
 DB 70 QVLTNIWLSQWTHYLOWNVSEYGVKTVTFPDQIWKPDILLYNSADERFDATFTN 129
 QY 127 VVRSGGSLYPPGIFKSTCKMDIAWFPDQHCMDKFGSWTYDGNQLDLVLKDEAGD 186
 DB 130 VLNVSSGHQCYLPPGIFKSSCVIDVRWFPDQHCXKLFKGSYGSWSLDLQMQE---AD 186
 QY 187 LSDFTNGEYWLIGMPGKNNTTYACCPPEYVDVTFIMRRRTLYFPNLIIVPCULISS 246
 DB 187 ISGYIPNGEWDLVGIPGKRSEFYECCKEPYDPDVTFTVTRRTLYGLNLLIPCVLLISA 246
 QY 247 MALLGFTLPDSEKLTGLVTSLLTFLNVAETLPQVSDAIPLLGYFNCIMPMVAS 306
 DB 247 LALLVFLPADSGEKISLTGLTLLSLTVFMLLVAEIMPATSDSVPLIAQFYASTMIIVGL 306

QY 307 SVVLTVVVLNYHRTADIHEMPQWIKSVFLOLFWILMSRPGKKTIRKTIIMNTRMREL 366
 DB 307 SVVTVIVLQYHHHPDGGKPKWTRVILLNWCAWFLRMKRPGEKVKVPACQHKQRCSL 366
 QY 367 ELKERSKSLANVLDDIDDFRHGPPPPNPGTASTGNL-----GPGC 407
 DB 367 ASVEMSAVA-----PPP---ASNGNLLYIGFRLDGVHCVPTPDSGVVC 407
 QY 408 SIFRTDFRRSFVRPSTMEDV--GGGLGSHHREHLILRELOFITARMKKADEAEELISDW 465
 DB 408 G-----RMACSPTHDEHLHGQPPGPDGLAKILEEVRYIANRFRCDSEAVGSEW 460
 QY 466 KFAAMVVDRECLFVFTLTITIAVAVLLSAPHII 499
 DB 461 KFAACVVDRICLMAFVSFTIITIGILMSAPNFV 494
 RESULT 10
 AAB50012
 ID AAB50012 standard; protein; 502 AA.
 XX AC AAB50012;
 XX DT 14-MAR-2001 (first entry)
 XX DE Wild-type human alpha7 ligand gated ion channel.
 XX KW Human; alpha7 nicotinic acetylcholine gated ion channel;
 XX KW 5-hydroxytryptamine; 5-HT3; calcium ion conductance.
 XX OS Homo sapiens.
 XX PN WO200073431-A2.
 XX PD 07-DEC-2000.
 XX PF 25-MAY-2000; 2000WO-US011862.
 XX PR 27-MAY-1999; 99US-0136174P.
 XX PA (PHAA) PHARMACIA & UPJOHN CO.
 XX PI Groppi VE, Wolfe ML, Berkenpas MB;
 XX WPI; 2001-061524/07.
 XX DR N-PSDB; AAC90380.
 XX PT Special cell culture medium for treating cells and for inducing mammalian
 PT cell lines to conduct calcium ions, comprising specified concentrations
 PT of ions of sodium, calcium and potassium at specified pH.
 XX PS Disclosure; Page 61-63; 77pp; English.
 XX CC The present sequence is wild-type human alpha7 nicotinic acetylcholine
 CC gated ion channel. The human alpha7 ion channel was used in the
 CC construction of an alpha7/5-hydroxytryptamine (5-HT3) chimeric ligand
 CC gated ion channel (see AAC90382 and AAB50014). The alpha7/5-HT3 chimeric
 CC ion channel can be expressed by recombinant cells in the present
 CC invention, resulting in preferential calcium ion conductance by the cells
 XX SQ Sequence 502 AA;

Query Match 47.7%; Score 1258.5; DB 4; Length 502;
 Best Local Similarity 48.6%; Pred. No. 1.4e-122;
 Matches 250; Conservative 75; Mismatches 138; Indels 51; Gaps 7;
 QY 8 LALLA-ILPVSEOGHEKRLNALLANVTLPVPANSEPLEVFGTLQOIIVDDEKN 66
 DB 10 LALAASLLHVSLSQGEFQRKLYKELYNPLRPVANDSQPLTVYFVLSLQIMVDDEKN 69
 QY 67 QLLITNIWLSLEWNYDNLRWNDSEYGVKDLRITPNKLWKPDLVLMYNSADEGFGDTYQTN 126

Db 70 QVLTNINWQMSWTHYLONNVSEYGVKTVRFPDQGIWKPDIILYNSADERDFATHTN 129

QY 127 VVVRSGGSLYVPPGIFKSTCKMDIAWPFDDQHCMDKFGSWTYDGNQLDLVKDEAGD 186

Db 130 VLVNSSGHCOYLPPIGIFKSCYDVRWFFDVQHCCKLFGSGWSYGGSLDLQOE---AD 186

QY 187 LSDFTNGWYLIQMPGKKNITTYACCPYVDVTFIMIRRTLYFFNLIIVPCVLISS 246

Db 187 ISGYIPNGEWDLVGIPGKSERFYECCKEYPDVTFTVMRRRTLYYGLNLLIPCVLISA 246

QY 247 MALLGFTLPDPSGKLTGLVTLLSLTVFLNVAETLPQVSDAIPILGTYFNCIMFWAS 306

Db 247 LALLVFLPADSGEKISLIGITVLLSVFMLVAEIMPATSDSVPLIAQYFASTMIIVGL 306

QY 307 SVVLTVVVLYNHHRTADIHEMPQWIKSVFLOLWPLIRMSRPGKTKITRKTIMMTRMREL 366

Db 307 SVVTVIVLVQHHHDPDGGKMPKWTIRVILLNWCWFLRMPGDKVRPACQHKQRCSL 366

QY 367 ELKERSKSLANVLIDDDFRHGPPPNSTAGNL-----GPGC 407

Db 367 ASVMSAVA-----PPP---ASNGNLLYIGRGLDGVHCVPTPDGSGVVC 407

QY 408 SIFRTDFRSFVRPSTMEDV---GGGLGSHHRELHLILRELQFITARMKKADEAEELISDW 465

Db 408 G-----RMACSPHDEHLHGQPPGDPDLAKILEEVRYIANFRQDESEAVCSEW 460

QY 466 KFAAMVDRFCLFVFTFTIATVAVLLSAPHII 499

Db 461 KFAACVDRCLMAFSVFTIICTIGILMSAPNFV 494

RESULT 11

ID ABG70492 standard; protein; 502 AA.

XX AC ABG70492;

XX DT 06-DEC-2002 (first entry)

XX DE Human neuronal nicotinic acetylcholine receptor alpha 7 subunit.

XX KW Human; neuronal nicotinic acetylcholine receptor; nNACHr; receptor;

XX KW ion flux; alpha 7 subunit.

XX OS Homo sapiens.

XX PN US6440681-B1.

XX PD 27-AUG-2002.

XX PF 07-JUN-1995; 95US-00487596.

XX PR 03-APR-1990; 90US-00504455.

XX PR 30-NOV-1992; 92US-00938154.

XX PR 08-MAR-1993; 93US-00028031.

XX PR 08-NOV-1993; 93US-00149503.

XX PA (MERI) MERCK & CO INC.

XX PI Elliott KJ, Ellis SB, Harpold NM;

XX DR WPI; 2002-711528/77.

XX DR N-FSDB; ABS54875.

PT Identifying antagonists or agonists of human neuronal nicotinic

PT acetylcholine receptors, by contacting recombinant cells with test

PT compound, and measuring ion flux of cells or binding of compound to

PT nNACHr.

XX Claim 101; Col 59-64; 56pp; English.

XX The invention relates to a method for identifying compounds that are

XX antagonists or agonists of human neuronal nicotinic acetylcholine

CC receptors (nNACHrs), by contacting recombinant cells with a test compound

CC and measuring ion flux, the electrophysiological response of the cells or

CC binding of the test compound to the nNACHr. The recombinant cells are

CC produced by transfection with a nucleic acid encoding at least one human

CC nNACHr (alpha or beta) subunit, such that the cells express an nNACHr

CC comprising one human subunit encoded by the transfected nucleic acid.

CC This sequence represents the alpha 7 subunit of the human nNACHr

XX polypeptide

XX Sequence 502 AA;

QY Query Match 47.7%; Score 1258.5; DB 5; Length 502;

Db Best Local Similarity 48.6%; Pred. No. 1.4e-122;

QY Matches 250; Conservative 75; Mismatches 138; Indels 51; Gaps 7;

Db 8 LALLA-LLPVSQGPHEKELNALLANVTLRPPVANESEPLEVRGLTLOQIIVDEKN 66

QY 10 LALAAALLHVSILQGEFQRLYKELVKNYNPLRPVANDSQPLTVVPSLLQLIMDVDEKN 69

Db 67 QLLITNWLSEWNYNLRWNSDEYGGVKDLRITPNKLMKPDVLMVNSADEGDFGTQTN 126

QY 70 QVLTNINWQMSWTHYLONNVSEYGVKTVRFPDQGIWKPDIILYNSADERDFATHTN 129

Db 127 VVVRSGGSLYVPPGIFKSTCKMDIAWPFDDQHCMDKFGSWTYDGNQLDLVKDEAGD 186

QY 130 VLVNSSGHCOYLPPIGIFKSCYDVRWFFDVQHCCKLFGSGWSYGGSLDLQOE---AD 186

QY 187 LSDFTNGWYLIQMPGKKNITTYACCPYVDVTFIMIRRTLYFFNLIIVPCVLISS 246

Db 187 ISGYIPNGEWDLVGIPGKSERFYECCKEYPDVTFTVMRRRTLYYGLNLLIPCVLISA 246

QY 247 MALLGFTLPDPSGKLTGLVTLLSLTVFLNVAETLPQVSDAIPILGTYFNCIMFWAS 306

Db 247 LALLVFLPADSGEKISLIGITVLLSVFMLVAEIMPATSDSVPLIAQYFASTMIIVGL 306

QY 307 SVVLTVVVLYNHHRTADIHEMPQWIKSVFLOLWPLIRMSRPGKTKITRKTIMMTRMREL 366

Db 307 SVVTVIVLVQHHHDPDGGKMPKWTIRVILLNWCWFLRMPGDKVRPACQHKQRCSL 366

QY 367 ELKERSKSLANVLIDDDFRHGPPPNSTAGNL-----GPGC 407

Db 367 ASVMSAVA-----PPP---ASNGNLLYIGRGLDGVHCVPTPDGSGVVC 407

QY 408 SIFRTDFRSFVRPSTMEDV---GGGLGSHHRELHLILRELQFITARMKKADEAEELISDW 465

Db 408 G-----RMACSPHDEHLHGQPPGDPDLAKILEEVRYIANFRQDESEAVCSEW 460

QY 466 KFAAMVDRFCLFVFTFTIATVAVLLSAPHII 499

Db 461 KFAACVDRCLMAFSVFTIICTIGILMSAPNFV 494

RESULT 12

ID ABG82435 standard; protein; 502 AA.

XX AC ABG82435;

XX DT 22-JAN-2003 (first entry)

XX DE Human neuronal NACHr alpha7 subunit.

XX KW Human; neuronal; nicotinic acetylcholine receptor; NACHr; drug screening;

XX KW immunochemistry; NACHr alpha7 subunit; receptor.

XX OS Homo sapiens.

XX PN WO200259266-A2.

XX PD 01-AUG-2002.

XX PF 29-OCT-2001; 2001WO-US050985.

XX

[illegible]

Search completed: May 7, 2004, 11:35:49
Job time : 44.0968 secs

XX		14-AUG-2002; 2002WO-US025765.	
XX	PF		
XX	PR	14-AUG-2001; 2001US-0312147P.	
XX	PR	01-NOV-2001; 2001US-0346382P.	
XX	PR	26-NOV-2001; 2001US-0333347P.	
XX	PA	(GEOH) GEN HOSPITAL CORP.	
XX	PA	(FARB) BAYER AG.	
XX	PI	Woolf C, D'urso D, Befort K, Costigan M;	
XX	DR	WPI; 2003-268312/26.	
XX	DR	GENBANK; P36544.	
XX	PT	New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.	
XX	PS	Claim 1; Page; 1017pp; English.	
XX	CC	The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.	
XX	SQ	Sequence 502 AA;	
	Query Match	47.7%; Score 1258.5; DB 7; Length 502;	
	Best Local Similarity	48.6%; Pred. No. 1.4e-122;	
	Matches 250; Conservative	75; Mismatches 138; Indels 51; Gaps 7	
Qy	8	LALLA-LIPVSEOGPEHXKELLNALLANNYTLRPPVANSEPLEVEFGLTQLQIIDVDEN 66	
Dd	10	LALAAALLHVLSLQGFBQRKLXKELYNNYNPLRPVANDSQPTVTVFSLQLQIMDVDEKN 69	
Qy	67	QLLIITINILSLFNWDNYNLRWNDSEYGGVKDLRITPNKLWKPDVLVLYNSADEGDFGTQTN 126	
Dd	70	QVLITINILQMSWTDHYLQNVNVEYPGVKTVPDPGQIWKPDILLVNSADERDAFTHTN 129	
Qy	127	VVVRSGGCLYPPGIFKSTCKMDTANFPFDQHCMDKFGSYTDGNQLDLVLKDBAGD 186	
Dd	130	VLVNSSGHGCYLPPGIFKSSCYIDVRWFEPDQHCCLKFGWSYGWSLDLQMQE---AD 186	
Qy	187	LSDFITNGSEWYLLIGMPGKKNTITYACCPBPYVDVTFITMRRTLLYFFENLVPCVLLIS 246	
Dd	187	ISGYIPNEGDVLVIPGRKSERYECCKEPYDPDTFTVMRRRLTYGLNLLIPCULLISA 246	
Qy	247	MALLGFTLPDPDSGEKILTGLVTILLSITVFLNLVAETLPQVSDAIPLLGTGFNCMFVNAS 306	
Dd	247	LALLVFLIPADSGEKISGITVLLSLTVMLIVAEIMPATSDSPVLAIOAFYSTMIIVGL 306	

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2004, 06:45:49 ; Search time 5514.12 Seconds
(without alignments)
3938.047 Million cell updates/sec

Title: US-09-303-232-6

Perfect score: 2640

Sequence: 1 MAPMAALALLALPVSEQ.....LFTIATVALLSAPHIIQ 501

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USFio.spool/US09303232/runat_07052004_101110_23883/app_query.fasta_1.2261
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09303232.CCN 1_1 8225 @runat_07052004_101110_23883 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2640	100.0	3029	3	AF143847 Heliothis
2	2640	100.0	3109	6	E58348 Nucleic aci
3	2640	100.0	3109	6	AX009614 Sequence
4	2194	83.1	2023	3	AF321447 Drosophil
5	2181	82.6	2023	3	AF321446 Drosophil
6	2177	82.5	1665	3	DME554209
7	2177	82.5	1699	3	BT011147 Drosophil
8	2177	82.5	2023	3	AF321445
9	2176.5	82.4	2068	3	AF321448
10	2156.5	81.7	2110	3	AF321449
11	1803.5	68.3	3629	3	AF143846
12	1803.5	68.3	3700	6	AX009612
13	1803.5	68.3	3701	6	E58347
14	1801.5	68.2	1683	3	DME554210
15	1600.5	68.2	2907	3	AF272778
16	1671	63.3	2834	3	AY036613
17	1609	60.9	2886	6	E58346
18	1609	60.9	2886	6	AX009610
19	1260	47.7	2087	9	HSARA7A
20	1258.5	47.7	1509	6	AX054567
21	1258.5	47.7	1509	9	HSNACHRA7
22	1258.5	47.7	1876	6	AR055255
23	1258.5	47.7	1876	6	AR071403
24	1258.5	47.7	1876	6	AR173187
25	1258.5	47.7	1876	6	AR224030
26	1258.5	47.7	1876	6	AR282833
27	1258.5	47.7	1876	6	AX719088
28	1258.5	47.7	1876	9	HSU62436
29	1258.5	47.7	2088	10	AF225980
30	1255.5	47.6	1848	10	MUSNARS
31	1255.5	47.6	2106	10	RATNARAD
32	1254.5	47.5	1590	6	AR261850
33	1254.5	47.5	1590	6	BD023656
34	1254.5	47.5	2107	9	AF486623
35	1253	47.5	2037	5	GGNNARA7A
36	1253	47.5	2696	5	GGA7NAREC
37	1252.5	47.4	1509	6	AX054575
38	1252.5	47.4	1916	9	AF385585
39	1252.5	47.4	1977	9	HSU40583
40	1250	47.3	1559	9	HUMA7NAR
41	1248.5	47.3	1509	6	AX054577
42	1246.5	47.2	3030	10	S53987
43	1245	47.2	1668	5	AY298752
44	1242.5	47.1	1509	6	AX054579
45	1240.5	47.0	1551	4	BTA7

ALIGNMENTS

RESULT 1

```

AF143847      3029 bp      mRNA      linear      INV 27-MAY-1999
LOCUS      Heliothis virescens putative nicotinic acetylcholine receptor alpha
DEFINITION      7-2 subunit mRNA, complete cds.
ACCESSION      AF143847
VERSION      AF143847.1
KEYWORDS      GI:4895006
SOURCE      Heliothis virescens (tobacco budworm)
ORGANISM      Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE      Schulte, T., Oellers, N. and Adamczewski, M.
AUTHORS      Putative alpha subunits of insect nicotinic acetylcholine receptors
TITLE      more similar to vertebrate alpha 7 subunits and C. elegans Ce21
than to other insect nicotinic acetylcholine receptor alpha
subunits
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 3029)
AUTHORS      Schulte, T., Oellers, N. and Adamczewski, M.
TITLE      Direct Submission
JOURNAL      Submitted (19-APR-1999) ZF-BTB, Bayer AG, Bldg. Q 18., Leverkusen
51368, Germany
FEATURES             Location/Qualifiers
     source           1..3029
                     /organism="Heliothis virescens"
                     /mol_type="mRNA"
                     /db_xref="taxon:7102"
                     95..1600
                     /note="hvnachra7-2"
                     /codon_start=1
                     /product="putative nicotinic acetylcholine receptor alpha
                     7-2 subunit"
                     /protein_id="AAD32698.1"
                     /db_xref="GI:4895007"
                     /translation="MAPMLAALALLIPVSEQGPHEKRLINALLANNVTLERPVANE
                     SEPLEVRFGLTQOI IDVDEKNQLITNIWLEWNDVNLRWNDSEYGVKDLRIITPN
                     KLVKPDVNLNSADEGDTQYTNVVRSGSCLYVPGIPEKSTCKMDIWFPPDDOH
                     CDMKFGWTYDGNQDILVLDKDEAGDLSDFITINGEWYLI GMPGKNITITYACCPYV
                     DVFTIMIRRTLYFNLIVPCVLISSMALLGFTLPDPSGEKLTGLVITLLSLTVFL
                     NLVAETLPQSDAIPGLTGTFNCIMFVASSVLTIVVLYNHYHRTADIHENPMWIKSV
                     FLOWLPLKMSRPGKIKTRITIMNTRMRELEKRSKSLLANLVIDDDFRHGPP
                     PPNSTASTGNLGPSCSIFRTDFRSFVRPTMEDVGGGLGSHHRELHLIRELQFIITA
                     RMKKADEAEALISDPKFAAMVDRFCULFVFTLITIAVAVLLSAPHIIVQ"
ORIGIN
Alignment Scores:
Pred. No.:      7,52e-245      Length:      3029
Score:          2640.00      Matches:      501
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              3          Gaps:      0

US-09-303-232-6 (1-501) x AF143847 (1-3029)

Qy      1 MetAlaProMetLeuAlaAlaLeuAlaLeuAlaLeuProValSerGluGlnGly 20
Db      95 ATGGCCCCATATGTTGGCGCCCTGGCGCTGCTGGCTTTCGCGCTATCGGAGCAAGGT 154

Qy      21 ProHisGluLysArgLeuLeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgPro 40
Db      155 CCTACGAGAGAGACTCTCGACCGCTTGTGGCGAACTACACACCCCTGGAGCGACCG 214

Qy      41 ValAlaAsnGluSerGluProLeuGluValArgPheGlyLeuThrLeuGlnIleLeu 60
Db      215 GTGGCCAAACGACGAGCAACCGGTAGAGGTCAGGTCAGGTCGGCTTCACCTTCAGCAATCAT 274

Qy      61 AspValAspGluLysAsnGlnLeuLeuIleThrAsnIleTyrLeuSerLeuGluTrpAsn 80
Db      275 GACGTGGGAGGAGAAATCACTACTATTAACCAATATATGGCTGCTGTTGGTGGGAAT 334

```

```

81  AspTyrAsnLeuArgTyrAsnAspSerGluTyrGlyValLysAspLeuArgIleThr 100
Db      335 GACTACAACTGAGTGGTGAACGACGAGTATGGCGGGTCAAGGACCTCAGGATCAGC 394

Qy      101 ProAsnLysLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp 120
Db      395 CCCAACAAAGTTGTGAAGCGGACGCTCTTATGTATATAATAGTGTCTGACGAGGGTTTTC 454

Qy      121 GlyThrTyrGlnThrAsnValValArgSerGlyGlySerCysLeuTyrValProPro 140
Db      455 GGGACCTTACCAGACCAACGCTGTGTGAGAGCGCGGCGAGTTCCTGTGTAGTCCACCT 514

Qy      141 GlyIlePheLysSerThrCysLysMetAspIleAlaTyrPhePheAspAspGlnHis 160
Db      515 GGCATATTCAGAGACCATGCAAGATGACATCGCGTGGTTCCTTCGACGACCAACAC 574

Qy      161 CysAspMetLysPheGlySerTyrTrpThrTyrAspGlyAsnGlnLeuAspLeuValLeuLys 180
Db      575 TGTGATATGAAGTTCGGTAGCTGGACATATGACGGCAATCAGTTGGATCTGGTGTCTAAAA 634

Qy      181 AspGluAlaGlyGlyAspLeuSerAspPheIleThrAsnGlyGluTyrTyrLeuIleGly 200
Db      635 GATGAGCGACGCGCGGATCTATCGGACTTCATAACAATGGGAGTGTATCTTAATAGGA 694

Qy      201 MetProGlyLysLysAsnThrIleThrTyrAlaCysCysProGluProTyrValAspVal 220
Db      695 ATGCCAGGCAAAAGAACACAAATAACATACGCGTGCTGCCCGAGCCCTACGTGACGATC 754

Qy      221 ThrPheThrIleMetIleArgArgArgThrLeuTyrTyrPhePheAsnLeuIleValPro 240
Db      755 ACCTTCACCATCATGATAAGAGACGAACTTGTACTACTTCTTCAACTCATGTCGTCGG 814

Qy      241 CysValLeuIleSerSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlu 260
Db      815 TGCGTGCTGATCTCATCGATGGCACTCTCGGCTTCACACTGCCACACGACTCCGGAGAG 874

Qy      261 LysLeuThrLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAla 280
Db      875 AAACCTCACACTTGGAGTCACATATTCTTCTATCGCTGACGGTGTTCCTCAACTGGTAGCC 934

Qy      281 GluThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIle 300
Db      935 GAGACCTCGCCACAGGCTCCGACGCTATCCCCCTGTTAGGAGCGTACTTCAATTCGATC 994

Qy      301 MetPheMetValAlaSerSerValValLeuThrValValLeuAsnTyrHisArg 320
Db      995 ATGTTTCATGTTAGCTGCTGCTGTGTACTGCTGTGTGGTACTCAATACCACTACGATCGA 1054

Qy      321 ThrAlaAspIleHisGluMetProGlnTrpIleLysSerValPheLeuGlnTrpLeuPro 340
Db      1055 ACAGCTGATATACATGAAATGCCACAGTGGATAAAATCAGTATTCTTCAATGGTGGCA 1114

Qy      341 TrpIleLeuArgMetSerArgProGlyLysIleThrArgLysThrIleMetMetAsn 360
Db      1115 TGGATACTCGCAATGTGAGGCCAGGGAAGAGATCACCGAAGAACTATTAATGATGAAC 1174

Qy      361 ThrArgMetArgGluLeuGluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnVal 380
Db      1175 ACGAGGATGAGGAGCTGGAACTGAAGAGAGAGTCTCGAAGTCTTCTGCTGGCGATGTT 1234

Qy      381 LeuAspIleAspAspPheArgHisGlyProProProProAsnSerThrAlaSerThr 400
Db      1235 CTAGATATTGATGATGACTTCAGACACGCGCTCCGCTCTCTAAACAGTACTGCCTCGACC 1294

Qy      401 GlyAsnLeuGlyProGlyCysSerIlePheArgThrAspPheArgArgSerPheValArg 420
Db      1295 GGGAAATTTGGACCTGGGTGCTCAATATTCGCGACGGAATTCGCGTGTGCTGTCGCT 1354

Qy      421 ProSerThrMetGluAspValGlyGlyLeuGlySerHisHisArgGluLeuHisLeu 440
Db      1355 CCGTCCAGCATGGAGACCTGGCGGGCGGTGGGTGACCACTCCGAGCTGCACCTC 1414

Qy      441 IleLeuArgGluLeuGlnPheIleThrAlaArgMetLysLysAlaAspGluAlaGlu 460

```

Db 1415 ATACTGAGAGAGTGCAGTTTCATACCGCCAGGATGAGAGCGCTGATGAGGAGCCGAG 1474
 QY 461 LeuIleSerAspTrpLysPheAlaAlaMetValValAspArgPheCysLeuPheValPhe 480
 Db 1475 CTGATCAGCGACTGGAAGTTTGCTGCGATGGTTGTCATAGGTTTTCCTGCTGCTTC 1534
 QY 481 ThrLeuPheThrIleAlaThrValAlaValLeuSerAlaProHisIleIleVal 500
 Db 1535 ACATTTTCAATCATCGGACAGTAGCTGCTGTTATCGGCACCGCATATCATCGTG 1594
 QY 501 Gln 501
 Db 1595 CAA 1597

RESULT 2

E58348
 LOCUS 3109 bp DNA linear PAT 18-JUN-2001
 DEFINITION Nucleic acid encoding insect acetyl choline receptor subunit.
 E58348
 ACCESSION
 VERSION E58348.1 GI:13019347
 KEYWORDS JP 2000023680-A/3.
 SOURCE Heliothis virescens (tobacco budworm)
 ORGANISM Heliothis virescens
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 Noctuoidea; Noctuidae; Heliothinae; Heliothis.
 1 (bases 1 to 3109)
 Martin, A., Nadja, E. and Thomas, S.
 Nucleic acid encoding insect acetyl choline receptor subunit
 Patent: JP 2000023680-A 3 25-JAN-2000;
 BAYER AG

COMMENT

OS Heliothis virescens
 PN JP 2000023680-A/3
 PD 25-JAN-2000
 PF 26-APR-1999 JP 1999118159
 PR 04-MAY-1998 DE 19819829.9
 PI MARTIN ADAMUTSUEUSUKI, NADJA ERASU, THOMAS SCHULTE PC
 C12N15/09, A01K67/033, C07K14/705, C07K16/28, C12N1/21, C12N5/10, PC
 C12Q1/68,
 PC G01N33/15, G01N33/50// (C12N1/21, C12R1/19), C12N15/00, C12N5/00 CC

FEATURES
source

FH Key Location/Qualifiers
 FT CDS (95). (1597).
 1..3109
 /organism="Heliothis virescens"
 /mol_type="genomic DNA"
 /db_xref="taxon:7102"

ORIGIN

Alignment Scores:
 Pred. No.: 7,79e-245 Length: 3109
 Score: 2640.00 Matches: 501
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-303-232-6 (1-501) x E58348 (1-3109)

QY 1 MetAlaProMetLeuAlaAlaLeuAlaLeuAlaLeuProValSerGluGlnGly 20
 Db 95 ATGGCCCCATATGTGGCGGCTTGGCGCTGCTGGCTTTGCTGCCGTATCGGAGCAAGGT 154
 QY 21 ProHisGluLysArgLeuLeuAsnAlaLeuAlaAsnTyrAsnThrLeuGluArgPro 40
 Db 155 CCTCAGAGAGAGACTCTCTGACGCGTTGCTGCGCACTACACACCTTGGAGCGACCG 214
 QY 41 ValAlaAsnGluSerGluProLeuGluValArgPheCysLeuThrLeuGlnGlnIle 60
 Db 215 GTGGCCAACAGAGACGACCGCTAGAGGTTCAGGTTCGGCTTGCCTTGCAGCAAAATCAIT 274

QY 61 AspValAspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsn 80
 Db 275 GAGCTGGACGAGAGAATCAACTACTTATAACCAATATATGGCTGTGCTGGAGTGAAT 334
 QY 81 AspTyrAsnLeuArgTrpAsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThr 100
 Db 335 GACTACACCTGAGGTGGAAACGACGAGGATATGGCGGGTCAAGGACCTCAGGATCACG 394
 QY 101 ProAsnLysLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp 120
 Db 395 CCCAACAAAGTTGTGGAAGCGGACGCTCTTATGTATAATAGTGTCTGACGAGGGTTTGAC 454
 QY 121 GlyThrTyrGlnThrAsnValValArgSerGlyGlySerCysLeuTyrValProPro 140
 Db 455 GGGACCTTACCAGACCAACGCTGGTGGTGAAGCGGCGGAGTTGCCCTGTACGTGCCACCT 514
 QY 141 GlyIlePheLysSerThrCysLysMetAspIleAlaTrpPheProPheAspAspGlnHis 160
 Db 515 GGCATATTCAAGAGCACATGCAAGATGGACATCGCTGGTTCCTTCGACGACCAACAC 574
 QY 161 CysAspMetLysPheGlySerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLys 180
 Db 575 TGTGATATGAAGTTTCGTAGCTGCACATATGACGGCAATCAGTTGGATCTGTGCTAAA 634
 QY 181 AspGluAlaGlyGlyAspLeuSerAspPheIleThrAsnGlyGluTrpTyrLeuIleGly 200
 Db 635 GATGAGCGAGCGCGCATCTATCGGACTTCATACAAATGGGAGTGGTATCTAATAGGA 694
 QY 201 MetProGlyLysLysAsnThrIleThrTyrAlaCysCysProGluProTyrValAspVal 220
 Db 695 ATGCCAGGCAAAAGAACACATAATAGCGGTGCTGCCCGGAGCCCTACGTGGAGCGTC 754
 QY 221 ThrPheThrIleMetIleArgArgThrLeuTyrTyrPhePheAsnLeuIleValPro 240
 Db 755 ACCTTCACCATCATGATAAGAGACGAACCTTGTACTACTTCTTCAACCTGATCGTCCG 814
 QY 241 CysValLeuIleSerSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlu 260
 Db 815 TGGTGTCTGATCTCATCGATGGCACTCTCGGCTTCACTGCCACGAGCTCCGGAGAG 874
 QY 261 LysLeuThrLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAla 280
 Db 875 AAACACACACTGGAGTCACTATCTCTATCGCTACGGGTGTTCTCAACCTGGTAGCC 934
 QY 281 GluThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIle 300
 Db 935 GAGACCTCGCCACAGGTCTCCGACGCTATCCCCCTGTAGGGACGTACTCAATTGCATC 994
 QY 301 MetPheMetValAlaSerSerValValLeuThrValValLeuAsnTyrHisArg 320
 Db 995 ATGTTTCATGTAGCGTCTGTGGTACTGACTGTGGTGTACTCAATTACCCACCATCGA 1054
 QY 321 ThrAlaAspIleHisGluMetProGlnTrpIleLysSerValPheLeuGlnTrpLeuPro 340
 Db 1055 ACAGCTGATATACATGAATGCCAGATGGATGAATAAATCAGTATTCCTACATGGTGCCA 1114
 QY 341 TrpIleLeuArgMetSerArgProGlyLysLysIleThrArgLysThrIleMetMetAsn 360
 Db 1115 TGGATACTCGAATGTGAGGCGCAGGGAAGAGATCACCAGGAAGACTATAATGATGAAC 1174
 QY 361 ThrArgMetArgGluLeuLeuLysGluArgSerSerLysSerLeuLeuAlaAsnVal 380
 Db 1175 ACAGATGAGGAGGTGGAACCTGAAGGAGAGGTGCTCGAAGTCTTGTGCGCAATGTT 1234
 QY 381 LeuAspIleAspAspPheArgHisGlyProProProProAsnSerThrAlaSerThr 400
 Db 1235 CTAGATATTGATGATGATCTTCAGACACGCGCTCCGCTCTTAACAGTACTGCCTCGACC 1294
 QY 401 GlyAsnLeuGlyProGlyCysSerIlePheArgThrAspPheArgSerPheValArg 420
 Db 1295 GGGAAATTTGGGACCTGGGTGCTCAATATCCGACGAGTTTCCGTGGTCCGTCCGT 1354
 QY 421 ProSerThrMetGluAspValGlyGlyLeuGlySerHisArgGluLeuHisLeu 440

1355 CCGTCCACGATGGAAGACGTGGGCGGCGTGGGTAGCCACCATCGCAGCTGCACCTC 1414
 441 ILleuArgGluLeuGlnPheIleThrAlaArgMetIlyslsAlaaspGluGluAlaGlu 460
 1415 ATA²CTGAGAGAGTGTGAGTTCATCAGCGGCAGATGAAGAAGCTGATGAGGAAGCGGAG 1474
 461 LeuIleSerAspTrpIlyslsPheAlaAlaMetValValAspArgPheCysLeuPheValPhe 480
 1475 CTGATCAGCGACTGGAAAGTTTGTCTGCATGGTGTGTGATAGGTTTTCCTGTTCTGTTGTC 1534
 481 ThrLeuPheThrIleLeuAlaThrValAlaValLeuLeuSerAlaProHisIleIleVal 500
 1535 ACAC²TTTTCACATCATCGGCAGTAGTGTCTCTGTTATCGGCACCGCATATCATCGTG 1594
 501 Gln 501
 1595 CAA 1597
 RESULT 3
 AX009614
 LOCUS AX009614 3109 bp DNA linear PAT 06-SEP-2000
 DEFINITION Sequence 5 from Patent EP0962528.
 ACCESSION AX009614
 VERSION AX009614.1 GI:99968846
 KEYWORDS Heliothis virescens (tobacco budworm)
 SOURCE Heliothis virescens
 ORGANISM Heliothis virescens
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 Noctuoidea; Noctuidae; Heliothinae; Heliothis.
 1
 Adamczewski, M.D., Schulte, T.D. and Oellers, N.D.
 Nucleic acids encoding acetylcholin-receptor subunits from insects
 Patent: EP 0962528-A 5 08-DEC-1999;
 BAYER AG (DE)
 FEATURES
 source 1..3109
 /organism="Heliothis virescens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:7102"
 95..1600
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC07501.1"
 /db_xref="GI:99968847"
 /db_xref="REMITREMBL:CAC07501"
 /translation="MAPMLAALALLPVSEQPEKRLNALLANYNLTLPFVANE
 SEPLEVRFGTLQOIIDVDEKNQILLITNIWLSWENDYNLRWDSYGGVKDLRIPTN
 KLMKPVLMYNSADGDTGYOTNVVRSGGCLYVPGPIFKSTCKMDIAWFPDPDVH
 CKMPGYSYDNGDLVDLXDEAGDLDSPFITNGEWVLIGMPGKNITLYACCEPVV
 DVFTIMIRRTLYYFENLIVPCVLISSMALLGFTLPDPSGEKLTIGVTLLISLTVFL
 NLVAETLPQSDAIPLDGTYNCFMFWASSVLTVVVLYNHHTADIHEMPQWIKSV
 FLQWPLWLRMSFPKKITRTKTMNTRMRELEKRSKSLLANVLIDDDFRHGPV
 PPNSTAGNLGPGCSIFRTDFRRSFVRPSTMEDVGGGLSGHREHLILURELQFITA
 RMKKADEAEELISDWKFAAMVDRFCLFVFTLFTIIATVAVLLSAPHILVQ"
 ORIGIN
 Alignment Scores:
 Pred. No.: 7.79e-245 Length: 3109
 Score: 2640.00 Matches: 501
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-303-232-6 (1-501) x AX009614 (1-3109)
 Qy 1 MetAlaProMetLeuAlaAlaLeuAlaLeuAlaLeuProValSerGluGlnGly 20
 Db 95 ATGGCCCCATGTGTGGCGGCGCTTGGCGCTGTCTGCTTGTCTGCGCATCGGACAGGT 154
 Ov 21 ProHisGluLysArgLeuLeuAsnAlaLeuAlaLeuAlaLeuAsnThrLeuGluArgPro 40


```
Qy 196 TrpTyrLeuIleGlyMetProGlyLysIleValThrTyrTyrAlaCysCysProGlu 215
Db 985 TGGTACTTGCTGTCATCGCGGAAAGAAATACAGTAGTCTACGCTGCTGCCGAGAA 1044
Qy 216 ProTyrValAspValThrPheThrIleMetIleArgArgThrLeuTyrTyrPhePhe 235
Db 1045 CCATATGTCGATATCACCTTTACTATACAAATTCGTCGCCGTACATTATATTTTTC 1104
Qy 236 AsnLeuIleValProCysValLeuIleSerSerMetAlaLeuLeuGlyPheThrLeuPro 255
Db 1105 AATTATATCGTCGATGTCATCTCATCGATGCGCCCTACTGGGCTTCACATGCCG 1164
Qy 256 ProAspSerGlyGluLysLeuThrLeuGlyValThrIleLeuLeuSerLeuThrValPhe 275
Db 1165 CCGGATTCGGCGAGAACTGACGCTGGGAGTTACAATCTTCTATGCTCACAGTGT 1224
Qy 276 LeuAsnLeuValAlaGluThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThr 295
Db 1225 CTCACCTTGTAGCTGAGACATTCGCCCAAGTATCTGATGCAATCCCTTGTAGGCAC 1284
Qy 296 TyrPheAsnCysIleMetPheMetValAlaSerSerValLeuThrValValValLeu 315
Db 1285 TACTTCAATTGTCATCATGTCGCTCGCTCGGTGGTGGTGCAGACAGTAGTGGTGC 1344
Qy 316 AsnTyrHisArgThrAlaAspIleHisGluMetProGlnTrpIleLysSerValPhe 335
Db 1345 AACTACCAACCATCGACAGCGGACATTCACGAGATGCCCGTGATCAAGTCGTTTC 1404
Qy 336 LeuGlnTrpLeuProTrpIleLeuArgMetSerArgProGlyLysLysIleThrArgLys 355
Db 1405 CTACAATGGCTGCTGATCTTCGATGGCTCGACCCGGTGCAGAGATTACACGCAAA 1464
Qy 356 ThrIleMetMetAsnThrArgMetArgGluLeuLysGluArgSerSerLysSer 375
Db 1465 ACAATACTATTAAAGCAATCGCATGAAGAGCTGAGCTAAAGAGCGCTCTCCAAATCC 1524
Qy 376 LeuLeuAlaAsnValLeuAspIleAspAspPheArgHisGlyProProProProAsn 395
Db 1525 CTGCTGGCAATGCTCTCGACATCGACAGCACTTCGGCAC----- 1566
Qy 396 SerThrAlaSerThrGlyAsnLeuGlyProGlyCysSerIlePheArgThrAspPheArg 415
Db 1567 ---ACAATATCTGGCTCCAAACCGCCATTCGCTCGT----- 1602
Qy 416 ArgSerPheValArgProSerThrMetGluAspValGlyGlyLeuGlySerHisHis 435
Db 1603 GCCAGCTTCGGTCCGCCACACGCTGGAGGAGCATCACGCGCATCGCTGCAATCAC 1662
Qy 436 ArgGluLeuHisLeuIleLeuArgGluLeuGlnPheIleThrAlaArgMetLysLysAla 455
Db 1663 AAAGATCTTCACTTAATCTCAAGAAATTGCAATTTATTACGCGCGGATCGCAAGCT 1722
Qy 456 AspGluGluAlaGluLeuSerAspTrpLysPheAlaAlaMetValValAspArgPhe 475
Db 1723 GACGACGAAGCGGAATTGATCGGCATTGGAAGTTTCGGCAATGGTTGTGATAGATT 1782
Qy 476 CysLeuPheValPheThrLeuPheThrIleIleAlaThrValAlaValLeuLeuSerAla 495
Db 1783 TGTTTAATTGTTTTCACGCTCTTACGATTAATGCAACGGTTACGCTGCTCTCCGCT 1842
Qy 496 ProHisIleIleValGln 501
Db 1843 CCGCACATAATCGTGCAA 1860
RESULT 5
AF321446 2023 bp mRNA linear INV 29-APR-2002
LOCUS Drosophila melanogaster nicotinic acetylcholine receptor Dalphae
DEFINITION subunit variant type II (nAcRalpha-30D) mRNA, complete cds,
alternatively spliced.
ACCESSION AF321446
VERSION AF321446.1 GI:20152846
```

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS

Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.

TITLE

Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, Dalphae5, Dalphae6 and Dalphae7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing

Genetics 160 (4), 1519-1533 (2002)

JOURNAL

21969411

MEDLINE

11973307

REFERENCE

2 (bases 1 to 2023)

Grauso, M. and Sattelle, D.B.

AUTHORS

Direct Submission

TITLE

Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional

Genetics Unit, University of Oxford, South Park Road, Oxford OX1

30X, UK

FEATURES

source

1. .2023

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/chromosome="2"

/map="30D1"

/dev_stage="embryo"

1. .2023

/gene="nAcRalpha-30D"

/gene="nAcRalpha-30D"

/note="ion channel; neurotransmitter transmembrane

receptor; alternatively spliced; contains exons 3b and 8a"

/product="nicotinic acetylcholine receptor Dalphae6 subunit

variant type II"

/protein_id="AA013393.1"

/db_xref="GI:20152847"

/translation="MDSPLASLSLVLLIFLAIKESCOCGPEKLLNHLSTVNTL

ERPVAESEPLEVKGFLTQOIIDVDERKNQLIINLWLSLEWNYNLRWNETEGGVK

DLRITPKMKPVDVLYNSADEGPDGYHTNIVVKGSGCLYVPPGIFKSTCKMDITW

FPDDQCEKMGSTYDQGLDLVLSNEDGSLDSFITNGEWYLLAMPGKNTIVYA

CCPEYVDITFTIQRRTLYYFNLIVPCVLISSMALLGFTLPDSEKILTLGVITL

LSITVFNLVAESMPTSDAVPLIGTFNCFIMFVASSVLTIVVLNVRHTADITHM

PPATKSVLOWLPWILEMGRGPKITKTLILSNRKELEKERSKSLLANVLDDID

DFRHTISGQTATGSSAFGRPTTVEEHHTAIGCNHKLHLILKELQFITARMKADD

EAEILGDWKFAMVVDRLIVFTLTITATVTVLLSAPIIVQ"

793. .794

/gene="nAcRalpha-30D"

/note="results in glycine to asparagine substitution;

compared to the sequence deposited in GenBank Accession

Number AE003626"

/replace="aa"

794

/gene="nAcRalpha-30D"

/note="results in glycine to serine substitution; compared

to variant clone"

/replace="a"

846

/gene="nAcRalpha-30D"

/note="results in asparagine to serine substitution;

compared to the sequence deposited in GenBank Accession

Number AE003626 and in variant clone"

/replace="a"

1314

/gene="nAcRalpha-30D"

/note="compared to variant clone"

/replace="a"

1560

/gene="nAcRalpha-30D"

/note="compared to variant clone"


```

gene
1. .1665
/gene="nAcRalpha-30D"
CDS
44. .1528
/gene="nAcRalpha-30D"
/codon_start=1
/product="nicotinic acetylcholine receptor subunit
Dalp6a6"
/protein_id="CAD86935.1"
/db_xref="GI:29466435"
/translat="MDSQPLPASLSLFLVILLIPLAIKESQCPGKRLNHLHLLSYNTPL
ERPVAENPEKPEVKGFLTQLOI IDVDEKNQNLNLWLSLENDYNLRWNETYGGVK
DLRITENKLPEDVLMYNSADEGDFGTVHTSVVKKGGSCLYVPGFIKSTCKMDITW
PPDDQHCMEKSGWYDGNQLDLVLSDEGDLSLDRTINGEWYLLAMPCKONTIUYA
CCPEYVDITFTIQIRRLTIFYFFNLIVPCVLISSMALLGTLTPDPSGKLTGLVTVIL
LSLTVFLNLVAESMPTSDAVPLITGVNCFMFASSVVLTVVVLNHHRTADIHEM
PFWTKSVFQLWPMLRMGRPKRIKRTILLSNRMKDELKELKSSKSLIANLIDDIQD
DPRITISGSQTAGSAGSFGRPTVEEHDTAIGCNKHDLHLILKELOFITARMERKADD
EAEILGDMKFAAMVVDRECLIVFTLTITIAITVTVLLSAPHIIVQ"
44. .121
/gene="nAcRalpha-30D"
122. .1525
/gene="nAcRalpha-30D"
/product="nicotinic acetylcholine receptor subunit
Dalp6a6"
sig_peptide
mat_peptide

```

ORIGIN

Alignment Scores:		
Pred. No.:	1.94e-200	Length:
Score:	2177.00	Matches:
Percent Similarity:	88.54%	Conservative:
Best Local Similarity:	81.62%	Mismatches:
Query Match:	82.46%	Indels:
DB:	3	Gaps:
		3
US-09-303-232-6 (1-501)	x	DME554209 (1-1665)

KEYWORDS
SOURCE
ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

gene

CDS

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Score:
1,99e-200	Matches:	1699
2177.00	Mismatches:	413
Percent Similarity:	Conservative:	35
Best Local Similarity:	Mismatches:	38
Query Match:	Indels:	20
DB:	Gaps:	3

ELI_CDNA.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1699)
Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,
Champe,M., Chavez,C., Dorsett,V., Dresnek,D., Farfan,D., Frise,E.,
George,R., Gonzalez,M., Guarin,H., Kronmiller,B., Li,P., Liao,G.,
Miranda,A., Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
Patel,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M.
and Celniker,S.
Direct Submission
Submitted (19-DEC-2003) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unsplined precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our Web site
(http://fruitfly.berkeley.edu) or send email to
cdna@fruitfly.berkeley.edu.

Location/Qualifiers

1..1699
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/strain="y; cn bw sp"
/db_xref="taxon:7227"

<1..1699
/gene="nAcRalpha-30D"
/db_xref="FLYBASE:FBgn0032151"
<1..1528
/note="Longest ORF"
/codon_start=2
/product="GHI518p"
/protein_id="AAR82815.1"
/db_xref="GI:40216008"
/db_xref="FLYBASE:FBgn0032151"
/translation="PFNPRATFYKAKMDSPSLASLSLFLVLLIFLAIIKESCGPHE
KELLNHLLSTNYTLERPVANSEPEVEFGTILQIIDVENKNQLLIINLWLSLEWD
YNLRNNEYGVKDLRIPTPKMLKPDLVMYNSADEGDGYHTFSVVVKHGGSCLYVP
PGIFTCKMIDTPFDFOCEMKFGSWTYDGNLDLVLEDSDGLSDPFTINGEWY
LIAMPKKNTIIVACCPPEYVDITFTIQIRRTLYVFNNLIPVCVLISMAILGFTEP
PDNGSKLTGLVTILLSLVFNVAESMETTSDAVPLGTGYENCIMFEWASSSVLTVV
VSNYHHRTADJHEMPWIKSFWLQWLPIWMKGPGRKITFKTILLSNRKKEBLKER
SKSKLLANVLDDDFRHITIGSGTAIGSSAFGRPTVEEHHTHCHNKHDHLLIK
ELQFITARRMRKADEALIGDWKFAMVVDVRFLVTFITFIATVTVLSAPHIIQV
"

US-09-303-232-6 (1-501) x BT011147 (1-1699)

Qy 2 AlaProMetLeuAlaAlaLeuAlaLeu-----LeuAlaLeuLeuPro 15

Db 50 TCCCGCTGCAGCGTGCTGCTGTTTCCTGCTGTTTCCTGCTGATAATAAA 109

Qy 16 ValSerGluGlnGlyProHisGluLysArgLeuLeuAsnAlaLeuAlaAsnTrpAsn 35

Db 110 GAAAGCTGTCAAGGACTCATGAAGAAGCGCTGCTGAACCATCTGCTGCCACCTACAT 169

Qy 36 ThrLeuGluArgProValalaAsnGluSerGluProLeuGluValArgPheGlyLeuThr 55

Db 170 ACCTCGAGGACGCGCTGGCCAATGAATCGAGCCCCTGAGGTTAAGTTCGAGCTGAGC 229

Qy 56 LeuGlnGlnleileAspValaspGluLysAsnGlnLeuLeuleleThrAsnlelrrPeu 75

Db 230 CTGACGAGATCATCGCTGGGACGAGAAGATCAGCTTCTCATACGAATCTTTGGCTT 289

Qy 76 SerLeuGluTrpAsnAspTyrrAsnLeuArgTrpAsnAspSerGluTyrrGlyVallys 95

Db 290 TCCTTGAGTGAAGCACTACAATCTGCCTGGAATGAACGSAATACGCGGGCTCAAG 349

Qy 96 AspleuArglleThrProasnlylsLeuTrpLysProaspValleuMetTyrrasnSerAla 115

Db 350 GATCTACGAATCACGCCCAACAGCTGTGGAAGCGGACGCTGCTATGTATACACAGCGCG 409

Qy 116 AspGluGlyPheAspGlyThrTyrrGlnThrAsnValValArgSerGlyGlySerCys 135

Db 410 GATGAGGATTTCATGGCAGCTATCACACCAGCTTGTGCTCAACATGGCGGCGAGTTGT 469

Qy 136 LeuTyrrValProGlyllePheLysSerThrCysLysMetAspileAlaTrpPhePro 155

Db 470 CTGTACGTGCCCGCTGGTATCTTCAAGACACATGCAAGATGCATCAGCTGGTTCCTCA 529

Qy 156 PheAspAspGlnHisCysaspMetlylsPheGlySerTrpThryrAspGlyAsnGlnLeu 175

Db 530 TTTCATGACCAACATTCGAAAATGAAATTCGGTAGTTGGACTTACGATGGAATCAGTTG 589

Qy 176 AspLeuValLeuLysAspGluAlaGlyLysAspLeuSerAspPheleleThrAsnGlyGlu 195

Db 590 GATTTGGTTTTGAGTTCGGAAGATGGAGGGGATCTTTCGATTTTCATCAAAATGGCGAG 649

Qy 196 TrpTyrrLeulleelyMetProGlyLysLysAsnThrilleThrTyrrAlaCysCysProGlu 215

Db 650 TGGTACTTGTTCCTGCTGCGGGAAGAAATAAGATAGTCTACGCTGCTGCCACCAA 709

Qy 216 ProTyrrValaspValThrPheThrilleMetlleArgArgArgThrLeuTyrrPhePhe 235

Db 710 CCATATGTGATATACCTTTATATACAAATTCGTCGCGTACATATATATATTTTTC 769

Qy 236 AsnLeulleValProCysValleulleSerSerMetAlaLeuLeuGlyPheThrLeuPro 255

Db 770 AATTTAATTTGGCCATGTGCTTAATCTCATCGATGGCCCTTCTGGGCTTCACATGGCG 829

Qy 256 ProAspSerGlyclulysLeuThrLeuGlyValThrilleLeuLeuSerLeuThrValPhe 275

Db 830 CGGATTCGGCGGAGAAACTGACGCTGGCGGTAATACTACTACTATCATTAACAGTATTT 889

Qy 276 LeuAsnLeuValalagluThrleuProGlnValSerAspAlalleProleuLeuGlyThr 295

Db 890 CTAACCTTTGTGCGCGAGTCCATGTCGCAACGTCGATGCTGTTCTCTTATATAGGACC 949

Qy 296 TyrPheAsnCysileMetPheMetValAlaSerSerValValLeuValValLeu 315

Db 950 TACTTCATTTGATCATGTTTCATGTCGCTGCTGCTGGTGGTGGTGGTGGTGGTGGT 1009

Qy 316 AsnTyrrHisArgThrAlaAspileHisGluMetProGlnTrpIlleLysSerValPhe 335

Db 1010 AACTACCACTATCGACAGCGGACATTCACGAGATGCCACCGTGGATCAAGTCGGTTTC 1069

Qy 336 LeuGlnTrpLeuProTrpIlleLeuArgMetSerArgProGlyLysLysleleThrArglys 355

Db 1070 CTACATGCTCCCTGGATCTTTCGAATGGTTCGACCGCGTACGAGATTACACGAAA 1129


```
QY 56 LeuGlnGlnIleIleAspValaspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeu 75
|||
Db 565 CTGCAGCAGATCATCGACGTGGATGAAAGAAATCAGATTCTGACCACAAATCGTGGTTA 624
|||
QY 76 SerLeuGluTrpAsnAspTrpAsnLeuArgTrpAsnAspSerGluTrpGlyGlyVallys 95
|||
Db 625 AATTTCGAGTGGACGACTACATCTGCGCTGGAATGAACGGAATACGGCGGGTCAAG 684
|||
QY 96 AspLeuArgIleThrProAsnLysLeuTrpLysProAspValLeuMetTyrAsnSerAla 115
|||
Db 685 GATCTACGAAATCAGCCCAACAGCTGTGGAAGCCGACGTCTCATGTACACAGCGG 744
|||
QY 116 AspGluGlyPheAspGlyThrTyrGlnThrAsnValValArgSerGlyGlySerCys 135
|||
Db 745 GATGAGGAGATTCGATGCGCATCACCAACATTTGTGTCAACATACGGCAGTTGT 804
|||
QY 136 LeuTyrValProProGlyIlePheLysSerThrCysLysMetAspIleAlaTrpPhePro 155
|||
Db 805 CTGTACGTGCCCCCTGTGTATCTTCAGAGCAGATGCAAGATAGACATCAGCTGGTCCCA 864
|||
QY 156 PheAspAspGlnHisCysAspMetLysPheGlySerTrpThrTyrAspGlyAsnGlnLeu 175
|||
Db 865 TTTGATGACCAACATTCGGAATGAAATTCGGTAGTTGGACTTACGATGGAATCAGTTG 924
|||
QY 176 AspLeuValLeuLysAspGluAlaGlyAspLeuSerAspPheIleThrAsnGlyGlu 195
|||
Db 925 GATTTCGTTTGAATTCGGAAGATGGAGGGGATCTTTCCGATTTCATACAAATGGCGAG 984
|||
QY 196 TrpTyrLeuIleGlyMetProGlyLysLysAsnThrIleThrTyrAlaCysCysProGlu 215
|||
Db 985 TGGTACTTGTCTGCCATGCGCGGAAGAAATACGATAGTCTACGCGCTGCTGCCAGAA 1044
|||
QY 216 ProTyrValAspValThrPheThrIleMetIleArgArgThrLeuTyrTyrPhePhe 235
|||
Db 1045 CCATATGTCGATATCACCCTTACTATACAAATTCGTCGCGTACATTATATATTTTTC 1104
|||
QY 236 AsnLeuIleValProCysValLeuIleSerSerMetAlaLeuLeuGlyPheThrLeuPro 255
|||
Db 1105 AATTATATCGTGCATGTGTCTAATCTCATCGATGCGCCCTACTGGGCTTCACATTGCCG 1164
|||
QY 256 ProAspSerGlyGluLysLeuThrLeuGlyValThrIleLeuLeuSerLeuThrValPhe 275
|||
Db 1165 CCGGATTTCGGCGGAGAACTGACCTGGAGTTTCAATCTTCTATGCTCACAGTGTTC 1224
|||
QY 276 LeuAsnLeuValAlaGluThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThr 295
|||
Db 1225 CTCACCTTGTAGCTGAGACATTCGCCCAAGTATCTGATGCAATCCCTTGTAGGCACC 1284
|||
QY 296 TyrPheAsnCysIleMetPheMetValAlaSerSerValValLeuThrValValValLeu 315
|||
Db 1285 TACTTCAATTGCATCATGTTTCATGTCGCCCTCGTGGTGGTGCACAGTAGTGGTGTCTC 1344
|||
QY 316 AsnTyrHisHisArgThrAlaAspIleHisGluMetProGlnTrpIleLysSerValPhe 335
|||
Db 1345 AACTACACCATCGACAGCGGACATTCACAGATGCCACCGTGGATCAATCGTGGTTC 1404
|||
QY 336 LeuGlnTrpLeuProTrpIleLeuArgMetSerArgProGlyLysLysIleThrArgLys 355
|||
Db 1405 CTACAATGGCTGCCCTGGATCTTGGATGGTGGTGGACCGGTCGCAAGATTACACGCAAA 1464
|||
QY 356 ThrIleMetMetAsnThrArgMetArgGluLeuGluLysGluArgSerSerLysSer 375
|||
Db 1465 ACAATACTATTAAAGCAATCGCATGAAGGAGCTGAGAGCTAAAGGAGCGCTCTCCAAATCC 1524
|||
QY 376 LeuLeuAlaAsnValLeuAsnValIleAsnAspPheArgHisGlyProProProAsn 395
|||
Db 1525 CTGTGGCCCAATGTCCTGCATCGACACGACGACTTCCGGCAC- 1566
|||
QY 396 SerThrAlaSerThrGlyAsnLeuGlyProGlyCysSerIlePheArgThrAspPheArg 415
|||
Db 1567 ---ACAATATCTGGCTCCCAACCCGCAATGGCTCGTCG- 1602
|||
QY 416 ArgSerPheValArgProSerThrMetGluAspValGlyGlyLeuGlySerHisHis 435
```

```
Db 1603 GCAGCTTCGGTCGCCCCCAACAAGCTGGAGGAGCATCACACGCCCATCGTGCATCAC 1662
|||
QY 436 ArgGluLeuHisLeuLeuLeuArgGluLeuGlnPheIleThrAlaArgMetLysLysAla 455
|||
Db 1663 AAAGATCTTCATCTAATCTCAAGAAATGCAATTTATTACGCGCGGATGCGCAAGCT 1722
|||
QY 456 AspGluGluAlaGluLeuIleSerAspTrpLysPheAlaMetValValAspArgPhe 475
|||
Db 1723 GACGACGAGCGAAATGTATCGCGGATGGAAGTTCCGCGCAATGGTTGTGGATAGATT 1782
|||
QY 476 CysLeuPheValPheThrLeuPheThrIleIleAlaThrValAlaValLeuLeuSerAla 495
|||
Db 1783 TGTTTAATTTGTTTTCACGCTCTTCAGATTATTCGACGGTTACGCTGCTCTCGCT 1842
|||
QY 496 ProHisIleIleValGln 501
|||
Db 1843 CCGCACATAATCGTGCAA 1860
|||
RESULT 9
AF321448
LOCUS
DEFINITION
2068 bp mRNA linear INV 29-APR-2002
Drosophila melanogaster nicotinic acetylcholine receptor dalpha6
subunit variant type IV (nAcRalpha-30D) mRNA, complete cds,
alternatively spliced.
ACCESSION
AF321448
VERSION
AF321448.1 GI:20152850
KEYWORDS
Drosophila melanogaster (fruit fly)
SOURCE
Drosophila melanogaster
ORGANISM
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 2068)
Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.
Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
Dalpaa5, Dalpaa6 and Dalpaa7, in Drosophila melanogaster Identify a
New and Highly Conserved Target of Adenosine Deaminase Acting on
RNA-Mediated A-to-I Pre-mRNA Editing
Genetics 160 (4), 1519-1533 (2002)
JOURNAL
MEDLINE
21969411
PUBMED
11973307
REFERENCE
2 (bases 1 to 2068)
Grauso, M. and Sattelle, D.B.
Direct Submission
Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional
Genetics Unit, University of Oxford, South Park Road, Oxford OX1
30X, UK
FEATURES
Location/Qualifiers
1..2068
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="30D1"
/dev_stage="embryo"
1..2068
/gene="nAcRalpha-30D"
379..1908
/gene="nAcRalpha-30D"
/gene="ion channel; neurotransmitter transmembrane
receptor; alternatively spliced; contains exons 3a, 3b and
9b; contains two repeats of the functional domain loop D"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalpaa6 subunit
variant type IV"
/protein_id="AA13395.1"
/db_xref="GI:20152851"
/translaton="MDSPLPASLSLFLVLIPLAIKSCQCPHEKRLNHLSTVNTL
ERPVAESEPLEVFGTLTQI IDVDKQNLTTNWLNLDERKQNLITNLWLSLEWN
DYNLRWNETEYGVKDLRTITPNKLWPDVLMYNSADEGFDGTYHTNIIVKSGSLYV
PPGIFKSTCKMDITWFPDQCEMFTGTYDGNLDVLNSDGDGLDFTNGEW
YLLAMPCKNTIVACCPPEPYVDITFTTIQIRRTTYFFNLIVPCVLISSMALLGFTL
```



```

QY 461 LeuileSerAspTrpIysPheAlaAlaMetValValAspArgPheCysLeuPheValPhe 480
Db 1783 TTGATCGCGGATTGGGAAGTTCGGCGCAATGGTTGTGGATAGATTGTTTAATTGTTTC 1842

QY 481 ThrLeuPheThrIleAlaThrValAlaValLeuLeuSerAlaProHisIleIleVal 500
Db 1843 ACGCTCTTCAGATTATTGCAACGGTTACGGTGTCTCTCCGCTCGCACATAAATCGTG 1902

QY 501 Gln 501
Db 1903 CAA 1905

RESULT 10
AF321449 2110 bp mRNA linear INV 29-APR-2002
LOCUS Drosophila melanogaster nicotinic acetylcholine receptor Dalp6a6
DEFINITION subunit variant type V (nAcRalpha-30D) mRNA, complete cds,
alternatively spliced.
ACCESSION AF321449
VERSION AF321449.1 GI:20152852
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 2110)
Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.
Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
Dalp6a5, Dalp6a6 and Dalp6a7, in Drosophila melanogaster Identify a
New and Highly Conserved Target of Adenosine Deaminase Acting on
RNA-Mediated A-to-I Pre-mRNA Editing
Genetics 160 (4), 1519-1533 (2002)
21969411
PUBMED 11973307
REFERENCE 2 (bases 1 to 2110)
Grauso, M. and Sattelle, D.B.
Direct Submission
Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional
Genetics Unit, University of Oxford, South Park Road, Oxford OX1
3QX, UK

FEATURES
source
1..2110
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="30D1"
/dev_stage="embryo"
1..2110
/gene="nAcRalpha-30D"
379..1950
/gene="nAcRalpha-30D"
/feature="ion channel; neurotransmitter transmembrane
receptor; alternatively spliced; contains exons 3a, 8a and
8b; contains two repeats of the transmembrane TW2 region"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalp6a6 subunit
variant type V"
/protein_id="AA013396.1"
/db_xref="GI:20152853"
/translation="MDSEPLASLSLFLVLLIFLAILIKESQCPHEKRLNHLSTYNTIL
ERPVAESEPELVKFLTLQQLIDVEKNQLITTNAMLNLEWINDYLNKRNETRYGVK
DLRITPNKLPDVMYNSADEGFDGTYHTNVIVKHNGSCLYVPPGIFKSTCKMDITW
PFDDQHCMEKFGSTVGNQDLVNSDGDLSDFITNGEWYLLAMPKKNITIVA
CCPPYVDITITQIRRTLYFFNLIVPCVLISMLLGLTLPDQSKLTGLGVTL
LSLTVFLNLVAESPSTSDAVPLIGVTLLSLTVFLNLVAETLPQVSDAIELLGTEN
CIMEWASSVVLTVVLNHHRTADIEHPMPWIKSVFLQWLPILRMGRPGRKTRKT
ILLNRMKELEKRSKSLIANVLIDDDFRHTISGTAIGSSASGRPTVEEH
TAICGNKDLHLILKELOFITARMRKADDEAELIDGMKFAAMVVDVRFCLIVFTLFTII
ATVTVLLSAPHIIQV"
846
/gene="nAcRalpha-30D"

variation

```

```

/note="results in methionine to isoleucine substitution;
compared to the sequence deposited in GenBank Accession
Number AE003626"
/replacement="a"
1401
/gene="nAcRalpha-30D"
/note="compared to the sequence deposited in GenBank
Accession Number AE003626"
/replacement="c"
1647
/gene="nAcRalpha-30D"
/note="compared to the sequence deposited in GenBank
Accession Number AE003626"
/replacement="c"
1770
/gene="nAcRalpha-30D"
/note="compared to the sequence deposited in GenBank
Accession Number AE003626"
/replacement="c"

variation

variation

variation

ORIGIN
Alignment Scores:
Pred. No.: 2,55e-198 Length: 2110
Score: 2156.50 Matches: 415
Percent Similarity: 83.55% Conservative: 32
Best Local Similarity: 77.57% Mismatches: 39
Query Match: 81.69% Indels: 49
DB: 3 Gaps: 4
US-09-303-232-6 (1-501) x AF321449 (1-2110)
QY 2 AlaProMetLeuAlaAlaLeuAlaLeu-----LeuAlaLeuLeuPro 15
Db 385 TCCCCGCTGCCAGCGTCTGCTGCTTTGCTCTGTGTGATCTTTCTGGCGATTAATAA 444
QY 16 ValSerGluGlnGlyProHisGluLysArgLeuLeuAsnAlaLeuLeuLeuAsnTyrAsn 35
Db 445 GAAAGCTGTCAAGGACCTCATGAAAGCGCTGCTGAACCATCTGCTGTCCACCTACAAT 504
QY 36 ThrLeuGluArgProValAlaAsnGluSerGluProLeuGluValArgPheGlyLeuThr 55
Db 505 ACGCTGGAGCGACCGCTGGCAATGAATCGAGCCCTGGAGGTTAAGTTGCGACTGACG 564
QY 56 LeuGlnGlnIleLeuAspValAspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeu 75
Db 565 CTGCAGCAGATCATCGACGCTGGATGAAAGAAATCAGATTCGACCAATATGCGTGTGTTA 624
QY 76 SerLeuGluTrpAsnAspTyrAsnLeuArgTrpAsnAspSerGluTyrGlyValIys 95
Db 625 AATTGGAGTGGACGACTACAAATCTGGCTGGAAATGAAACGGAATACGGCGGGTCAAG 684
QY 96 AspLeuArgIleThrProAsnLysLeuTrpLysProAspValLeuMetTyrAsnSerAla 115
Db 685 GATCTAGCAATACAGCCCAACAGCTGTGGAAGCCCGACGCTCATGTACAAACAGCGCG 744
QY 116 AspGluGlyPheAspGlyThrTyrGlnThrAsnValValArgSerGlySerCys 135
Db 745 GATGAGGATTCGATGCGACGATATCACCCACAAATTTGTTGTTCAACATACCGGAGTTGT 804
QY 136 LeuTyrValProGlyIlePheLysSerThrCysLysMetAspIleAlaTrpPhePro 155
Db 805 CTGTACGTGCCCCCTGGTATCTTCAAGAGCACATGCAAGATGGACATCACGTGTTCCCA 864
QY 156 PheAspAspGlnHisCysAspMetLysPheGlySerTrpThrTyrAspGlyAsnGlnLeu 175
Db 865 TTTGATGACCAACATTCGAAATGAAATTCGGTAGTTGGAATTCAGATGGAATTCAGTTG 924
QY 176 AspLeuValLeuLysAspGluAlaGlyAspLeuSerAspPheIleThrAsnGlyGlu 195
Db 925 GATTGTGTTTGAATTCGAGATGGAGGGGATCTTCCCGATTTTCATACCAATATGGCGAG 984
QY 196 TrpTyrLeuIleGlyMetProGlyLysAsnThrIleThrTyrAlaCysCysProGlu 215

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDs
ORIGIN
Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
US-09-303-232-6 (1-501) x AF143846 (1-3629)
QY 2 AlaProMetLeuAlaAlaLeuAlaLeuAlaLeuLeuProValSerGluGln---Gly 20
Db 368 GGCCCGCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGG 427
QY 21 ProHisGluLysArgLeuLeuAlaAlaLeuLeuAlaAlaAsnTyrAsnThrLeuGluArgPro 40
Db 428 TACCACGAGAAGCGGCTACTGCACCACCTATTGGACCCTACCAACGCTACTGGAGAGGCC 487
QY 41 ValAlaAsnGluSerGluProLeuGluValArgPheGlyLeuThrLeuGlnGlnIleIle 60
Db 488 GTTCGGTCAACAGAGAGCGCCGCTGAGCTCTCCITTCGGCTCACGCTCATGCAGATCATC 547
QY 61 AspValAspGluLysAsnGlnLeuLeuThrAsnIleTyrLeuSerLeuGluTyrAsn 80
Db 548 GACGTGGACGAGAGAACAGCTTTTATAACAACATCTGGCTAAACTAGAGTGGAT 607
QY 81 AspTyrAsnLeuArgTyrAsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThr 100
Db 608 GATATGAATTCAGGTGGAACACTTCAGATTTCGGCGGGGTCAAAGATTAAAGAGTGCCA 667
QY 101 ProAsnLysLeuTyrLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp 120